

# CONVERGENCE RATES FOR HIERARCHICAL GIBBS SAMPLERS

OLIVER JOVANOVSKI\* AND NEAL MADRAS

*Department of Mathematics and Statistics, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3*

**ABSTRACT.** We establish some results for the rate of convergence in total variation of a Gibbs sampler to its equilibrium distribution. This sampler is motivated by a hierarchical Bayesian inference construction for a gamma random variable. Our results apply to a wide range of parameter values in the case that the hierarchical depth is 3 or 4, and are more restrictive for depth greater than 4. Our method involves showing a relationship between the total variation of two ordered copies of our chain and the maximum of the ratios of their respective co-ordinates. We construct auxiliary stochastic processes to show that this ratio does converge to 1 at a geometric rate.

**Key words and phrases:** Convergence Rate, Hierarchical Gibbs Sampler, Markov Chain, Coupling, Gamma Distribution, Stochastic Monotonicity

## 1. INTRODUCTION

There is a significant amount of theory showing that a Markov chain satisfying some fairly general conditions (see for example [1]) will converge to an equilibrium in distribution, as well as in the stronger measure of total variation. Mere knowledge of convergence is often not enough, and it may be of both theoretical and practical interest to consider the rate at which convergence proceeds. In particular, deriving an upper bound on the rate of convergence would provide a rigorously defined degree of certainty to how far this Markov chain is from its equilibrium distribution, and it would allow to assess the efficiency of an algorithm.

This has been our main objective in this paper, where the model in question is motivated by the following hierarchical Bayesian inference scenario: we are given a real number  $x > 0$  with the information that it was drawn from a  $\Gamma(a_1, u_1)$  distribution, defined by the probability density function

$$f(z) = \frac{u_1^{a_1}}{\Gamma(a_1)} z^{a_1-1} e^{-zu_1}$$

where  $a_1 > 0$  is given. The inverse scale parameter  $u_1$  is itself the product of a random sampling from an independent  $\Gamma(a_2, u_2)$  distribution. Once again it is assumed that  $a_2 > 0$  is a given constant, while  $u_2$  is sampled in an analogous manner. This process continues until we reach  $u_n \sim \Gamma(a_{n+1}, b)$ , where now both  $a_{n+1} > 0$  and  $b > 0$  are given. The joint density of  $(x, u_1, \dots, u_n)$  is therefore defined by

$$(1.1) \quad p(x, z_1, \dots, z_n) \propto x^{a_1-1} \left( \prod_{i=1}^n z_i^{a_i+a_{i+1}-1} \right) \exp \left( \sum_{i=1}^{n+1} -z_i z_{i-1} \right)$$

where for convenience we set  $z_0 := x$  and  $z_{n+1} := b$ . We conclude from (1.1) that for  $1 \leq i \leq n$ , the conditional distribution of  $u_i$  given everything else is

$$u_i | x, u_{j \neq i} \sim \Gamma(a_i + a_{i+1}, u_{i-1} + u_{i+1})$$

The resulting posterior distribution of  $(u_1, \dots, u_n)$  (i.e. given  $x$  as well as all other parameters) is therefore defined by the density function

$$(1.2) \quad g(z_1, \dots, z_n) \propto \left( \prod_{i=1}^n z_i^{a_i+a_{i+1}-1} \right) \exp \left( \sum_{i=1}^{n+1} -z_i z_{i-1} \right)$$

Should one wish to sample from it, however, it would be quite challenging to do so directly due to its complicated structure.

The Gibbs sampler [4] has been a very popular MCMC algorithm in obtaining a sample from a probability distribution that is otherwise difficult to work with. In its fundamental form, this algorithm works on a vector  $u$  by selecting (systematically, randomly or otherwise) one of the vector's components  $u_i$  and updating this component only, by drawing from the probability distribution of  $u_i$  given  $(u_{j \neq i})$ .

In this paper we define a variant of the Gibbs sampler which converges to the equilibrium distribution of our model. For  $n = 4$ , it can be characterised as a simultaneous updating algorithm: at each step in time, we set both components of  $u^{t+1}$  (after showing that the problem can be simplified to a 2 dimensional Markov chain) to be random with a distribution derived conditionally from  $u^t$ . For  $n > 4$ , we consider two different Markov chains given in (2.28) and (2.29). The evolution of both is elaborate, but we ultimately settle with (2.29) since it proves to be more favourable to our method.

General convergence results have been derived for some Gibbs samplers (e.g. [5]), however due to their limitations it is often not possible to infer quantitative bounds directly from these results. A frequently used method to derive such bounds is to couple two copies of a Markov chain (for a detailed discussion on this subject see [2]), and evaluate the likelihood of coalescence at some future time. It may also be quite useful to seek out an appropriate partial order on the state space, and attempt to couple in a stochastically monotone manner that preserves this partial order for all time (e.g. [3]). This has been our approach, with a few notable peculiarities: we define the partial order  $\preceq$  and consider initial vectors  $u^0 \preceq \tilde{v}^0$ . But rather than keeping track of the two Markov chains, we observe  $u^t$  in tandem with another stochastic process  $v^t$  that serves as a majorant (in the partial order  $\preceq$ ) to both copies of the Markov chain. We then provide an upper bound in Corollary 13 on the rate which the ratio  $v^t/u^t$  converges geometrically to 1, which is ultimately used to obtain an upper bound on the rate at which the two chains coalesce.

**1.1. The problem.** We will start by considering the following Markov chain on  $\mathbb{R}_+^n$ : for  $a_1, a_2, \dots, a_{n+1} \in \mathbb{R}_+$  and  $\gamma_i \sim \Gamma(a_i + a_{i+1}, 1)$  independent, we define the map

$$(1.3) \quad (u_1, u_2, \dots, u_n) \longrightarrow \left( \frac{\gamma_1}{x + u_2}, \frac{\gamma_2}{u_1 + u_3}, \dots, \frac{\gamma_n}{u_{n-1} + b} \right)$$

where  $x, b \in \mathbb{R}_+$ . An immediate consequence is that for  $t > 1$  we can eliminate the odd numbered co-ordinates by considering the 2-step (sub)Markov chain derived from (1.3). If  $n$  is odd this becomes

$$(u_2^{t+1}, u_4^{t+1}, \dots, u_{n-1}^{t+1}) = \left( \frac{\gamma_2^{t+1}}{\frac{\gamma_1^t}{x+u_2^{t-1}} + \frac{\gamma_3^t}{u_2^{t-1}+u_4^{t-1}}}, \frac{\gamma_4^{t+1}}{\frac{\gamma_3^t}{u_2^{t-1}+u_4^{t-1}} + \frac{\gamma_5^t}{u_4^{t-1}+u_6^{t-1}}}, \dots, \frac{\gamma_{n-1}^{t+1}}{\frac{\gamma_{n-2}^t}{u_{n-3}^{t-1}+u_{n-1}^{t-1}} + \frac{\gamma_n^t}{u_{n-1}^{t-1}+b}} \right)$$

and similarly, for  $n$  even this would only differ in the final co-ordinate,

$$(u_2^{t+1}, u_4^{t+1}, \dots, u_n^{t+1}) = \left( \frac{\gamma_2^{t+1}}{\frac{\gamma_1^t}{x+u_2^{t-1}} + \frac{\gamma_3^t}{u_2^{t-1}+u_4^{t-1}}}, \frac{\gamma_4^{t+1}}{\frac{\gamma_3^t}{u_2^{t-1}+u_4^{t-1}} + \frac{\gamma_5^t}{u_4^{t-1}+u_6^{t-1}}}, \dots, \frac{\gamma_n^{t+1}}{\frac{\gamma_{n-1}^t}{u_{n-2}^{t-1}+u_n^{t-1}} + b} \right)$$

This allows us to reduce the original problem on  $\mathbb{R}_+^n$  by considering the Markov chain defined by a random function on  $\mathbb{R}_+^{\lfloor \frac{n}{2} \rfloor}$  which for  $n$  odd is given by:

$$(1.4) \quad f_n(u_2, u_4, \dots, u_{n-1}) = \left( \frac{\gamma_2}{\frac{\gamma_1}{x+u_2} + \frac{\gamma_3}{u_2+u_4}}, \frac{\gamma_4}{\frac{\gamma_3}{u_2+u_4} + \frac{\gamma_5}{u_4+u_6}}, \dots, \frac{\gamma_{n-1}}{\frac{\gamma_{n-2}}{u_{n-3}+u_{n-1}} + \frac{\gamma_n}{u_{n-1}+b}} \right)$$

A similar definition results for the case where  $n$  is even. Note that if  $Y^t, U^t$  are instances of the Markov chains defined by (1.3) and (1.4) respectively, then  $U^t \sim \left( Y_2^{2t}, Y_4^{2t}, \dots, Y_{\lfloor \frac{n}{2} \rfloor}^{2t} \right)$ .

Another way of considering the process in (1.4) is to look at the Markov chain which sequentially updates its co-ordinates as follows: for  $i \in \{1, 2, \dots, n\}$  let

$$\bar{P}_i(v, w) := \left( \prod_{j \neq i} \delta_{v_j}(w_j) \right) h_i(w_i | v) dw_i$$

where  $h_i(w_i | v)$  is the density function of  $\frac{\gamma_i}{v_{i-1} + v_{i+1}}$  given  $v$ , and where for convenience we have defined  $v_0 := x$  and  $v_{n+1} := b$ . Let

$$(1.5) \quad \bar{P} := \bar{P}_1 \bar{P}_3 \cdots \bar{P}_{2\lceil \frac{n}{2} \rceil - 1} \bar{P}_2 \bar{P}_4 \cdots \bar{P}_{2\lfloor \frac{n}{2} \rfloor}$$

or in other words  $\bar{P}$  updates all odd-numbered co-ordinates first, followed by the even numbered co-ordinates. Observe that the transition kernel  $\bar{P}$  obeys the same law as the random function in (1.4),

$$\bar{P}(v, \cdot) \sim f_n(v_2, \dots, v_{2\lfloor \frac{n}{2} \rfloor})$$

Now let  $\pi$  be the probability measure on  $\mathbb{R}_+^n$  with density function (1.2). Then it is easy to conclude that  $\pi(dz) \bar{P}_i(z, dy) = \pi(dy) \bar{P}_i(y, dz)$ . So in particular this implies  $\pi \bar{P}_i = \pi$ , which in turn implies  $\pi \bar{P} = \pi$ . This shows that  $\pi$  must be the unique stationary distribution associated with the Markov chain in (1.4). From the previous remark on the relationship between (1.3) and (1.4), it follows that  $\pi$  is also the stationary distribution of (1.3).

We can now state our main results. For  $n = 4$ , let  $\mathcal{U}^t$  and  $\mathcal{V}^t$  be two copies of the Markov chain and let  $d_{TV}$  denote the total variation metric on probability measures of a probability space  $\Omega$ , defined by

$$d_{TV}(\nu_1, \nu_2) := \sup_{A \subseteq \Omega} |\nu_1(A) - \nu_2(A)|$$

For two random variables  $X \sim \nu_1$  and  $Y \sim \nu_2$ , we let  $d_{TV}(X, Y) := d_{TV}(\nu_1, \nu_2)$ . We will also define the condition

$$(1.6) \quad a_1 + a_4 > 1, a_2 + a_5 > 1, a_2 + a_3 > 1, a_3 + a_4 > 1$$

Then

**Theorem 1.** *Suppose that (1.6) holds. If  $J_0 \leq \eta$ , then  $\exists r < 1$  such that for  $t \geq d'$ ,  $d_{TV}(\mathcal{U}^{t+3}, \mathcal{V}^{t+3}) \leq r^{\lfloor \frac{t}{d'} \rfloor} (1 + 3(a_2 + a_3 + a_4 + a_5)(R_0 - 1))$ . For general values of  $J_0$ , we have that  $d_{TV}(\mathcal{U}^{t+3}, \mathcal{V}^{t+3}) \leq r^{\lfloor \frac{t}{2d'} \rfloor} (1 + 3(a_2 + a_3 + a_4 + a_5)(R_0 - 1)) + \frac{\max\{J_0, \eta\}}{\eta} \beta^{\lfloor \frac{t}{2} \rfloor + 3}$  for some  $\beta < 1$ .*

Here  $J_0$  and  $R_0$  depend on the initial conditions  $\mathcal{U}^0$  and  $\mathcal{V}^0$ , while all other terms are constants to be defined later. Our need for condition (1.6) becomes clear in section 2.1.2. If we let  $\mathcal{U}^0 = (1, 1, 1, 1)$  and  $\mathcal{V}^0 \sim \pi$  then

$$\textbf{Corollary 2.} \quad d_{TV}(\mathcal{U}^{t+3}, \pi) \leq \mathbb{E}_\pi[R_0] r^{\lfloor \frac{t}{2d'} \rfloor} (1 + 3(a_2 + a_3 + a_4 + a_5)) + \left( \frac{\mathbb{E}_\pi[J_0]}{\eta} + 1 \right) \beta^{\lfloor \frac{t}{2} \rfloor + 3}$$

All terms on the right-hand side of the last inequality are constants, and the quantities  $\mathbb{E}_\pi[R_0]$  and  $\mathbb{E}_\pi[J_0]$  depend only on  $\pi$  and can be estimated easily. For  $n = 2m > 4$  (we make an observation later that the analysis for odd  $n$  is nearly identical), we let  $\mathcal{U}^t$  and  $\mathcal{V}^t$  evolve according to (2.29). Then

**Theorem 3.** *Suppose that  $\max\{\zeta_2, \dots, \zeta_{2m}, \xi_2, \dots, \xi_{2m}\} < 1$ . Then*

$$d_{TV}(\mathcal{U}^{t+3}, \mathcal{V}^{t+3}) \leq r^{\lfloor \frac{t}{2d'} \rfloor} (1 + 3(a_2 + a_3 + \dots + a_{2m+1})(R_0 - 1)) + \frac{\max\{J_0, \eta\}}{\eta} \beta^{\lfloor \frac{t}{2} \rfloor + 3}$$

Here the terms  $\zeta_2, \dots, \zeta_{2m}$  and  $\xi_2, \dots, \xi_{2m}$  are defined in (2.41) and (2.44) respectively, and depend only on the parameters  $x, b$  and  $\{a_i\}$ .

**1.2. Outline of our proof.** Essentially the proof of Theorem 1 and Theorem 3 are quite similar, both relying on a coupling argument. In Section 1.3 we define a partial order ' $\preceq$ ' on  $\mathbb{R}_+^{\lfloor \frac{n}{2} \rfloor}$  and show that we can couple two copies  $\{u^t, v^t\}$  of (1.4), with the initial condition  $u^0 \preceq v^0$ , in a stochastically monotone manner, thus preserving the order  $u^t \preceq v^t$  for all time  $t$ . In the beginning of Section 2 we show that if  $R_t$  is a process that serves as an upper bound for the ratio  $\max_i \left\{ \frac{v_i^t}{u_i^t} \right\}$ , then the rate of convergence of  $R_t \rightarrow 1$  can be related to the rate at which (1.4) converges to equilibrium. Therefore, our focus becomes the defining of such a process and showing that it converges to 1 at a geometric rate.

We define  $v^t$  in Section 2.1 (for the case  $n = 4$ ) to be a stochastic process adapted to the same filtration as  $u^t$ , with the property that it is an upper bound to (in the sense of  $\preceq$ ) a faithful copy of (1.4) started at  $v^0$ . This allows us to define  $R_t$  and has the additional quality of being strictly monotone decreasing. This alone does not guarantee that  $R_t \rightarrow 1$  quickly (or at any pace, for that matter). But the rate at which  $R_t$  approaches 1 does depend on the size of the values  $u_2^t$  and  $u_4^t$ , and we show that if often enough these two values are neither too large nor too small, then  $R_t \rightarrow 1$  at a geometric rate. To fulfill this condition, we define a number of auxiliary processes in Section 2.1.1 (and show their existence in Section 2.1.2) that serve as an upper bound for the terms  $\left\{ u_2^t, u_4^t, \frac{1}{u_2^t}, \frac{1}{u_4^t} \right\}$ , and we show that they are frequently bounded from above by some constant  $\eta$ .

The case  $n > 4$  is treated in Section 2.3. We define a Markov chain somewhat different from (1.4), for the purpose of obtaining a monotone decreasing process  $R_t$  that has the desired properties mentioned above. The proof of Theorem 3 follows in an analogous manner to what we have for Theorem 1, however finding the required auxiliary processes proves to be more elusive. We show their existence under certain constraints on the parameters.

**1.3. Stochastically monotone coupling.** For  $u, v \in \mathbb{R}_+^n$ , define the partial order  $u \preceq v$  to mean  $u_i \leq v_i$  for even  $i$ , and  $u_i \geq v_i$  for odd  $i$ . For the 'reduced' chain (1.4) on  $\mathbb{R}_+^{\lfloor \frac{n}{2} \rfloor}$  we can take this partial order to imply the same (since we are only concerned with the even co-ordinates, this would mean that we have pointwise inequality in the same direction at every co-ordinate).

Suppose we couple two copies of (1.3),  $u^0 \preceq v^0$ , by employing the same random variables  $\{\gamma_i^t\}$  in both copies (we will refer to this as the 'uniform coupling'). Then  $u^t \preceq v^t$  for all times  $t$ . Therefore if we couple in this manner two copies commencing at some arbitrary initial points  $\mathcal{U}^0, \mathcal{V}^0 \in \mathbb{R}_+^n$ , we can take  $m = \min \{\mathcal{U}_1^0, \dots, \mathcal{U}_n^0, \mathcal{V}_1^0, \dots, \mathcal{V}_n^0\}$  and  $M = \max \{\mathcal{U}_1^0, \dots, \mathcal{U}_n^0, \mathcal{V}_1^0, \dots, \mathcal{V}_n^0\}$ , and define

$$\begin{aligned} v^0 &:= (m, M, m, \dots) \in \mathbb{R}_+^n \\ u^0 &:= (M, m, M, \dots) \in \mathbb{R}_+^n \end{aligned} \quad (1.7)$$

i.e. we are setting  $v_{2j+1}^0 = u_{2j+2}^0 = m$  and  $v_{2j+2}^0 = u_{2j+1}^0 = M$ . And by observing that  $u^0 \preceq \{\mathcal{U}^0, \mathcal{V}^0\} \preceq v^0$ , we conclude that  $\mathcal{U}^t$  and  $\mathcal{V}^t$  are perpetually 'squeezed' between  $u^t$  and  $v^t$  (or in other words  $u^t \preceq \{\mathcal{U}^t, \mathcal{V}^t\} \preceq v^t$ ). We can justify with Corollary 5 why it suffices to consider the coupled pair  $(u^t, v^t)$  in order to bound  $d_{TV}(\mathcal{U}^t, \mathcal{V}^t)$ .

**Lemma 4.** Suppose that  $0 < \beta_1 < \beta_2 < \beta_3 < \beta_4$ , and let  $z_i \sim \Gamma(\alpha, \beta_i)$  Then  $d_{TV}(z_2, z_3) \leq d_{TV}(z_1, z_4)$

*Proof.* Let  $f_1, f_2, f_3, f_4$  be the respective density functions. By the following property of total variation (see Theorem 5.7 of [7])

$$d_{TV}(z_i, z_j) = 1 - \int \min(f_i(y), f_j(y)) dy$$

it is enough to show that  $\min\{f_1(y), f_4(y)\} \leq \min\{f_2(y), f_3(y)\}$  for all  $y$ . Note first that for  $i, j \in \{1, 2, 3, 4\}$  with  $i < j$ ,

$$\begin{aligned} f_i(y) &\geq f_j(y) \\ \iff \beta_i^\alpha \exp(-\beta_i y) &\geq \beta_j^\alpha \exp(-\beta_j y) \\ \iff y &\geq \frac{\alpha(\ln(\beta_j) - \ln(\beta_i))}{\beta_j - \beta_i} \end{aligned} \quad (1.8)$$

Let

$$g(\beta, \kappa) := \frac{\alpha(\ln(\beta) - \ln(\kappa))}{\beta - \kappa}$$

then

$$\frac{\partial g}{\partial \kappa} = \frac{\alpha(1 - \frac{\beta}{\kappa} + \ln(\frac{\beta}{\kappa}))}{(\beta - \kappa)^2}$$

The numerator of this equation is non-positive for all  $\beta, \kappa \in \mathbb{R}_+$ . This can be seen by observing that the function  $\ln(z) - z$  achieves a global maximum on  $(0, \infty)$  at  $z = 1$  with value  $-1$ . Hence  $1 - \frac{\beta}{\kappa} + \ln(\frac{\beta}{\kappa}) \leq 0$ , and since  $g(\beta, \kappa)$  is non-increasing in  $\kappa$ , we have the following relation:

$$(1.9) \quad g(\beta_4, \beta_3) \leq g(\beta_4, \beta_2) \leq g(\beta_4, \beta_1) = g(\beta_1, \beta_4) \leq g(\beta_1, \beta_3) \leq g(\beta_1, \beta_2)$$

Then from (1.8) and (1.9) it follows that

$$\begin{aligned} f_1(y) &\leq \min\{f_2(y), f_3(y)\} \text{ on } [0, g(\beta_1, \beta_3)] \\ f_4(y) &\leq \min\{f_2(y), f_3(y)\} \text{ on } [g(\beta_4, \beta_2), \infty) \end{aligned}$$

hence

$$\min(f_1(y), f_4(y)) \leq \min\{f_2(y), f_3(y)\} \text{ on } [0, g(\beta_1, \beta_3)] \cup [g(\beta_4, \beta_2), \infty) = [0, \infty)$$

□

We apply the uniform coupling until time  $t$ , and given this outcome we couple  $(\mathcal{U}_{|\mathcal{U}^t}^{t+1}, \mathcal{V}_{|\mathcal{V}^t}^{t+1}, u_{|u^t}^{t+1}, v_{|v^t}^{t+1})$  in the following “one-shot” manner (described in [8] in further detail): for each co-ordinate  $i$ , we take  $\mathcal{U}_{i,|\mathcal{U}^t}^{t+1}$  to be the x-coordinate of a uniformly chosen point from the area under the graph of the density function  $f_{\mathcal{U}_i}$  of  $\mathcal{U}_{i,|\mathcal{U}^t}^{t+1}$ , and set  $\mathcal{V}_{i,|\mathcal{V}^t}^{t+1} = \mathcal{U}_{i,|\mathcal{U}^t}^{t+1}$  if this point also lies below the graph of the density function  $f_{\mathcal{V}_i}$  of  $\mathcal{V}_{i,|\mathcal{V}^t}^{t+1}$ . Otherwise take  $\mathcal{V}_{i,|\mathcal{V}^t}^{t+1}$  to be the x-coordinate of a uniformly chosen point from the area below the graph of  $f_{\mathcal{V}_i}$  and above the graph of  $f_{\mathcal{U}_i}$ . Similarly if the point  $\mathcal{U}_{i,|\mathcal{U}^t}^{t+1}$  is also below the graph of the density function  $f_{u_i}$  of  $u_{i,|u^t}^{t+1}$ , we set  $u_{i,|u^t}^{t+1} = \mathcal{U}_{i,|\mathcal{U}^t}^{t+1}$ , and otherwise we chose a point uniformly from the the area below the graph of  $f_{u_i}$  and above the graph of  $f_{\mathcal{U}_i}$ . Finally, if also  $u_{i,|u^t}^{t+1}$  is below the graph of  $f_{v_i}$ , we set  $v_{i,|v^t}^{t+1} = u_{i,|u^t}^{t+1}$ . Otherwise choose an independent point uniformly from the area that is below the graph of  $f_{v_i}$  and above the graph of  $f_{u_i}$ , and set  $v_{i,|v^t}^{t+1}$  to be the x-coordinate of this point. It is easy to verify that this is indeed a coupling of  $(\mathcal{U}_{|\mathcal{U}^t}^{t+1}, \mathcal{V}_{|\mathcal{V}^t}^{t+1}, u_{|u^t}^{t+1}, v_{|v^t}^{t+1})$ .

**Corollary 5.** *With one-shot coupling at time  $t + 1$ , we have  $d_{TV}(\mathcal{U}^{t+1}, \mathcal{V}^{t+1}) \leq \mathbb{P}[u^{t+1} \neq v^{t+1}]$*

*Proof.* We will first observe the following

$$\begin{aligned} d_{TV}(\mathcal{U}_{|\mathcal{U}^t}^{t+1}, \mathcal{V}_{|\mathcal{V}^t}^{t+1}) &\leq \mathbb{P}[\mathcal{U}_{|\mathcal{U}^t}^{t+1} \neq \mathcal{V}_{|\mathcal{V}^t}^{t+1}] \\ &= \mathbb{P}[\cup_i \{\mathcal{U}_{i,|\mathcal{U}^t}^{t+1} \neq \mathcal{V}_{i,|\mathcal{V}^t}^{t+1}\}] \\ &\leq \mathbb{P}[\cup_i \{u_{i,|u^t}^{t+1} \neq v_{i,|v^t}^{t+1}\}] \\ &= \mathbb{P}[u_{|u^t}^{t+1} \neq v_{|v^t}^{t+1}] \end{aligned}$$

The inequality in the third line is a consequence of the previous lemma and the fact that  $\min\{f_{\mathcal{U}_i}, f_{\mathcal{V}_i}\} \geq \min\{f_{u_i}, f_{v_i}\}$  implies  $\{\mathcal{U}_{i,|\mathcal{U}^{t-1}}^t \neq \mathcal{V}_{i,|\mathcal{V}^{t-1}}^t\} \subseteq \{u_{i,|u^{t-1}}^t \neq v_{i,|v^{t-1}}^t\}$ . This inequality holds for any outcome of  $(\mathcal{U}^t, \mathcal{V}^t, u^t, v^t)$  so long as the partial order  $u^t \preceq \{\mathcal{U}^t, \mathcal{V}^t\} \preceq v^t$  persists. But by the previous remarks this is always the case for the uniform coupling, hence the statement of the corollary. □

## 2. THE RATIO $R_t$

From here on we will mainly consider the ‘reduced’ Markov chain defined by (1.4), however at times it will be useful to refer to odd numbered co-ordinates as derived from the transition kernel  $\bar{P}$  in (1.5).

We assume in this section that  $u^t = \left(u_2^t, u_4^t, \dots, u_{2\lfloor \frac{n}{2} \rfloor}^t\right) \preceq \left(v_2^t, v_4^t, \dots, v_{2\lfloor \frac{n}{2} \rfloor}^t\right) = v^t$ , so that  $\frac{v_i^t}{u_i^t} \geq 1$ . Let  $R_t$  be a non-increasing  $\mathcal{F}_t$ -measurable process such that  $R_t \geq \max_i \left\{ \frac{v_i^t}{u_i^t} \right\}$ . Then  $u^t = v^t$  if  $R_t = 1$ . The uniform coupling defined for (1.3) in the previous section can also be easily applied to (1.4) as well as the result of Corollary 5 (where in the final step at time  $t+1$  one would couple  $\left(\gamma_1^{t+1}, \gamma_3^{t+1}, \dots, \gamma_{2\lfloor \frac{n}{2} \rfloor - 1}^{t+1}\right)$  uniformly, while taking the 'one shot' approach for  $\left(\gamma_2^{t+1}, \gamma_4^{t+1}, \dots, \gamma_{2\lfloor \frac{n}{2} \rfloor}^{t+1}\right)$  as described in the section preceding Corollary 5). Under these assumptions we can now relate  $R_t$  to  $\mathbb{P}[u^{t+1} \neq v^{t+1} | \mathcal{F}_t]$  where  $\mathcal{F}_t := \sigma(\gamma_1^1, \dots, \gamma_n^1, \dots, \gamma_1^t, \dots, \gamma_n^t)$ .

**Lemma 6.** *Applying one-shot coupling at time  $t+1$ , we have*

$$\mathbb{P}[u^{t+1} \neq v^{t+1} | \mathcal{F}_t] \leq 1 - R_t^{-\sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} (a_{2i} + a_{2i+1})}$$

*Proof.* Let  $h_{u_{2i}}(y | \mathcal{F}_t, \gamma_{2i-1}^{t+1}, \gamma_{2i+1}^{t+1})$  and  $h_{v_{2i}}(y | \mathcal{F}_t, \gamma_{2i-1}^{t+1}, \gamma_{2i+1}^{t+1})$  be the conditional density functions of  $u_{2i}^{t+1}$  and  $v_{2i}^{t+1}$  given  $\mathcal{F}_t, \gamma_{2i-1}^{t+1}, \gamma_{2i+1}^{t+1}$  (and therefore, also given  $u_{2i-1}^{t+1}, u_{2i+1}^{t+1}, v_{2i-1}^{t+1}, v_{2i+1}^{t+1}$  where  $u_{2i-1}^{t+1} = \gamma_{2i-1}^{t+1} / (u_{2i-2}^t + u_{2i-2}^t)$  and we define  $u_{2i+1}^{t+1}, v_{2i-1}^{t+1}$  and  $v_{2i+1}^{t+1}$  similarly). These represent gamma random variables with shape parameters given by  $a_{2i} + a_{2i+1}$  and scale parameters  $u_{2i-1}^{t+1} + u_{2i+1}^{t+1}$  and  $v_{2i-1}^{t+1} + v_{2i+1}^{t+1}$  respectively, as can be seen from the definition of the transition kernel  $\bar{P}$ . Then

$$h_{u_{2i}}(y | \mathcal{F}_t, \gamma_{2i-1}^{t+1}, \gamma_{2i+1}^{t+1}) \geq \left( \frac{u_{2i-1}^{t+1} + u_{2i+1}^{t+1}}{v_{2i-1}^{t+1} + v_{2i+1}^{t+1}} \right)^{a_{2i} + a_{2i+1}} h_{v_{2i}}(y | \mathcal{F}_t, \gamma_{2i-1}^{t+1}, \gamma_{2i+1}^{t+1})$$

hence

$$\min \{h_{u_{2i}}(y | \mathcal{F}_t, \gamma_{2i-1}^{t+1}, \gamma_{2i+1}^{t+1}), h_{v_{2i}}(y | \mathcal{F}_t, \gamma_{2i-1}^{t+1}, \gamma_{2i+1}^{t+1})\} \geq \left( \frac{u_{2i-1}^{t+1} + u_{2i+1}^{t+1}}{v_{2i-1}^{t+1} + v_{2i+1}^{t+1}} \right)^{a_{2i} + a_{2i+1}} h_{v_{2i}}(y | \mathcal{F}_t, \gamma_{2i-1}^{t+1}, \gamma_{2i+1}^{t+1})$$

As described in the lead-up to Corollary 5, we take  $u_{2i}^{t+1}$  to be the x-coordinate of a uniformly chosen point from the area under the graph of  $h_{u_{2i}}(y | \mathcal{F}_t, \gamma_{2i-1}^{t+1}, \gamma_{2i+1}^{t+1})$ , and set  $v_{2i}^{t+1} = u_{2i}^{t+1}$  if this point is also below the graph  $h_{v_{2i}}(y | \mathcal{F}_t, \gamma_{2i-1}^{t+1}, \gamma_{2i+1}^{t+1})$ . Otherwise take  $v_{2i}^{t+1}$  to be the x-coordinate of a uniformly chosen point from the area below the graph of  $h_{v_{2i}}(y | \mathcal{F}_t, \gamma_{2i-1}^{t+1}, \gamma_{2i+1}^{t+1})$  and above the graph of  $h_{u_{2i}}(y | \mathcal{F}_t, \gamma_{2i-1}^{t+1}, \gamma_{2i+1}^{t+1})$ . The result is that

$$\begin{aligned} \mathbb{P}[u_{2i}^{t+1} \neq v_{2i}^{t+1} | \mathcal{F}_t, \gamma_{2i-1}^{t+1}, \gamma_{2i+1}^{t+1}] &= 1 - \int \min \{h_{u_{2i}}(y | \mathcal{F}_t, \gamma_{2i-1}^{t+1}, \gamma_{2i+1}^{t+1}), h_{v_{2i}}(y | \mathcal{F}_t, \gamma_{2i-1}^{t+1}, \gamma_{2i+1}^{t+1})\} \\ &\leq 1 - \left( \frac{u_{2i-1}^{t+1} + u_{2i+1}^{t+1}}{v_{2i-1}^{t+1} + v_{2i+1}^{t+1}} \right)^{a_{2i} + a_{2i+1}} \\ &\leq 1 - R_{t+1}^{-a_{2i} - a_{2i+1}} \\ &\leq 1 - R_t^{-a_{2i} - a_{2i+1}} \end{aligned}$$

where  $R_{t+1}$  is the process defined above and derived under the hypothetical continuation of the uniform coupling. Since the last inequality is independent of  $\gamma_{2i-1}^{t+1}, \gamma_{2i+1}^{t+1}$ , we also get  $\mathbb{P}[u_{2i}^{t+1} \neq v_{2i}^{t+1} | \mathcal{F}_t] \leq 1 - R_t^{-a_{2i}-a_{2i+1}}$ . Therefore

$$\begin{aligned} \mathbb{P}\left[u^{t+1} \neq v^{t+1} \mid \mathcal{F}_t, \gamma_1^{t+1}, \gamma_3^{t+1}, \dots, \gamma_{2\lceil \frac{n}{2} \rceil - 1}^{t+1}\right] &= \mathbb{P}\left[\cup_i \{u_{2i}^{t+1} \neq v_{2i}^{t+1}\} \mid \mathcal{F}_t, \gamma_1^{t+1}, \gamma_3^{t+1}, \dots, \gamma_{2\lceil \frac{n}{2} \rceil - 1}^{t+1}\right] \\ &= 1 - \prod_i \mathbb{P}\left[\{u_{2i}^{t+1} = v_{2i}^{t+1}\} \mid \mathcal{F}_t, \gamma_1^{t+1}, \gamma_3^{t+1}, \dots, \gamma_{2\lceil \frac{n}{2} \rceil - 1}^{t+1}\right] \\ &\leq 1 - \prod_i R_t^{-a_{2i}-a_{2i+1}} \\ &= 1 - R_t^{-\sum_i (a_{2i}+a_{2i+1})} \end{aligned}$$

and once again since  $R_t^{-\sum_i (a_{2i}+a_{2i+1})}$  is independent of  $(\gamma_1^{t+1}, \gamma_3^{t+1}, \dots, \gamma_{2\lceil \frac{n}{2} \rceil - 1}^{t+1})$ , we get the desired result.  $\square$

**2.1. Special case  $n = 4$ .** In this section we will consider the Markov chain defined by the random functions  $\{f_n^t\}$  in (1.4) for the case where  $n = 4$ , and with initial conditions  $u^0 \preceq v^0$  that have the additional property  $\frac{u_2^0}{u_4^0} = \frac{v_2^0}{v_4^0}$ . Note from (1.7) that this may already be assumed. We proceed by defining the Markov chain  $u^t$ , and from here on we let  $v^t$  be the process defined in the following manner: set

$$(u_2^{t+1}, u_4^{t+1}) = f_4^{t+1}(u_2^t, u_4^t) = (f_{4,1}^{t+1}(u_2^t, u_4^t), f_{4,2}^{t+1}(u_2^t, u_4^t))$$

and let

$$(\tilde{v}_2^{t+1}, \tilde{v}_4^{t+1}) = f_4^{t+1}(v_2^t, v_4^t) = (f_{4,1}^{t+1}(v_2^t, v_4^t), f_{4,2}^{t+1}(v_2^t, v_4^t))$$

Then define

$$(2.1) \quad (v_2^{t+1}, v_4^{t+1}) := \max\left\{\frac{\tilde{v}_2^{t+1}}{u_2^{t+1}}, \frac{\tilde{v}_4^{t+1}}{u_4^{t+1}}\right\} (u_2^{t+1}, u_4^{t+1})$$

Observe that:

1. The equality of the ratios is always preserved:  $\frac{u_2^{t+1}}{u_4^{t+1}} = \frac{v_2^{t+1}}{v_4^{t+1}}$ .
2.  $(\tilde{v}_2^{t+1}, \tilde{v}_4^{t+1}) \preceq (v_2^{t+1}, v_4^{t+1})$ , hence by monotonicity the process  $v^t$  is always “greater than or equal to” a copy of the Markov chain started at  $v^0$  and coupled uniformly with  $u^t$ .
3. From the ratio above we also get  $\frac{v_2^{t+1}}{u_2^{t+1}} = \frac{v_4^{t+1}}{u_4^{t+1}}$ .

Hence if we let  $R_t := \frac{v_2^t}{u_2^t} = \frac{v_4^t}{u_4^t}$ , then

$$\begin{aligned} (2.2) \quad \frac{v_2^{t+1}}{u_2^{t+1}} &= \max\left\{\frac{\tilde{v}_2^{t+1}}{u_2^{t+1}}, \frac{\tilde{v}_4^{t+1}}{u_4^{t+1}}\right\} \\ &= \max\left\{\left(\frac{\frac{\gamma_3^{t+1}}{u_2^t + u_4^t} + b}{\frac{\gamma_3^{t+1}}{v_2^t + v_4^t} + b}\right), \left(\frac{\frac{\gamma_3^{t+1}}{u_2^t + u_4^t} + \frac{\gamma_1^{t+1}}{u_2^t + x}}{\frac{\gamma_3^{t+1}}{v_2^t + v_4^t} + \frac{\gamma_1^{t+1}}{v_2^t + x}}\right)\right\} \\ &= \frac{v_2^t}{u_2^t} \cdot \max\left\{\left(\frac{\frac{\gamma_3^{t+1}}{u_2^t} + bu_4^t}{\frac{\gamma_3^{t+1}}{u_2^t} + bv_4^t}\right), \left(\frac{\frac{\gamma_3^{t+1}}{u_2^t} + \frac{\gamma_1^{t+1}}{1 + \frac{x}{u_2^t}}}{\frac{\gamma_3^{t+1}}{u_2^t} + \frac{\gamma_1^{t+1}}{1 + \frac{x}{v_2^t}}}\right)\right\} \\ &\leq \frac{v_2^t}{u_2^t} Q_t \end{aligned}$$

where

$$Q_t = \max \left\{ \left( \frac{\gamma_3^{t+1} + \frac{\gamma_1^{t+1}}{1 + \frac{x}{v_2^t}}}{\gamma_3^{t+1} + \frac{\gamma_1^{t+1}}{1 + \frac{x}{v_2^t}}} \right), \left( \frac{\gamma_3^{t+1} + bv_4^t}{\gamma_3^{t+1} + bv_4^t} \right) \right\} \leq 1$$

The inequality in (2.2) is justified by

**Lemma 7.** *Suppose that  $0 < a < b$ . Then  $g(x, y) := (\frac{x}{b} + y)/(\frac{x}{a} + y)$  is decreasing in  $x$  and increasing in  $y$ , for all  $x, y > 0$ .*

*Proof.* Follows from calculus. □

Therefore

$$\mathbb{E}[R_{t+1}] \leq R_0 \mathbb{E} \left[ \prod_{j=0}^t Q_j \right]$$

It can be easily observed that the ratio  $R_t$  satisfies the condition stated in the paragraph preceding Lemma 6. The aim now is to obtain from the previous inequality an expression of the form

$$\mathbb{E}[R_{t+1}] \leq 1 + C_{R_0} \prod_{j=1}^{t+1} r_j$$

where  $r_j < 1$  and  $r_j$  is frequently bounded from above by some  $r < 1$  (the exact meaning of this will become apparent following the definition of  $\bar{S}_t$  in (2.11)). Note that in order to achieve this, it suffices to have for all  $t \geq 0$

$$(2.3) \quad \mathbb{E}[Q_t R_t] \leq r_{t+1} (\mathbb{E}[R_t] - 1) + 1$$

Let  $\mathcal{F}_t := \sigma((\gamma_1^1, \dots, \gamma_4^1), \dots, (\gamma_1^t, \dots, \gamma_4^t))$ . We can consider (2.3) by conditioning on this filtration

$$(2.4) \quad \mathbb{E}[Q_t R_t] = \mathbb{E}[R_t \mathbb{E}[Q_t \mid \mathcal{F}_t]]$$

and we may approximate  $\mathbb{E}[Q_t \mid \mathcal{F}_t]$  with the aid of the following lemmas.

Let  $\mu_1 = \mathbb{E}[\gamma_3] = a_3 + a_4$  and  $\mu_2 = \mathbb{E}[\gamma_1 - \frac{1}{3}] = a_1 + a_2 - \frac{1}{3}$ .

**Lemma 8.** *Let  $S$  be a  $\mathcal{F}_t$ -measurable stopping time. Then  $\mathbb{E}[Q_S R_S] \leq \mathbb{E}[\dot{r}_S (R_S - 1)] + 1$  ,*

where  $\dot{r}_t = 1 - 1/\max \left\{ \left( \frac{4\mu_1}{\mu_2} + 4 \right) \left( \frac{u^t}{x} + \frac{x}{v_2^t} + 2 \right), 4 + \frac{4\mu_1}{bv_4^t} \right\}$ .



*Proof.* By [6] we have  $\mathbb{P}[\gamma_3 \leq \mu_1] \geq \frac{1}{2}$  and  $\mathbb{P}[\gamma_1 \geq \mu_2] \geq \frac{1}{2}$  and so by the previous lemma

$$\begin{aligned}
\mathbb{E}[Q_S | \mathcal{F}_S] &\leq \frac{1}{4} \cdot \max \left\{ \left( \frac{\mu_1 + \frac{\mu_2}{1 + \frac{x}{v_2^S}}}{\mu_1 + \frac{\mu_2}{1 + \frac{x}{v_2^S}}} \right), \left( \frac{\mu_1 + bu_4^S}{\mu_1 + bv_4^S} \right) \right\} + 1 \cdot \frac{3}{4} \\
&= \frac{1}{4} \cdot \max \left\{ \left( 1 - \frac{\frac{1}{1 + \frac{x}{v_2^S}} - \frac{1}{1 + \frac{x}{u_2^S}}}{\frac{\mu_1}{\mu_2} + \frac{1}{1 + \frac{x}{v_2^S}}} \right), \left( 1 - \frac{bv_4^S - bu_4^S}{\mu_1 + bv_4^S} \right) \right\} + 1 \cdot \frac{3}{4} \\
&\leq \frac{1}{4} \cdot \max \left\{ 1 - \frac{\left(1 - \frac{1}{R_S}\right)}{\left(\frac{\mu_1}{\mu_2} + 1\right) \left(1 + \frac{x}{v_2^S}\right) \left(1 + \frac{x}{u_2^S}\right) \frac{v_2^S}{x}}, 1 - \frac{\left(1 - \frac{1}{R_S}\right)}{1 + \frac{\mu_1}{bv_4^S}} \right\} + 1 \cdot \frac{3}{4} \\
&\leq 1 - \frac{\left(1 - \frac{1}{R_S}\right)}{\max \left\{ \left(\frac{4\mu_1}{\mu_2} + 4\right) \left(\frac{u_2^S}{x} + \frac{x}{v_2^S} + 2\right), 4 + \frac{4\mu_1}{bv_4^S} \right\}} \\
(2.5) \quad &= \dot{r}_S + \frac{1 - \dot{r}_S}{R_S}
\end{aligned}$$

Substituting this into (2.4), we get the desired result.  $\square$

The next step will be to show that often enough  $\dot{r}_t \leq r$  for some  $r < 1$ , which by the inequality (2.5) would result in an expression of the form given by (2.3).

**2.1.1. Super-martingale-type auxiliary processes.** The results in this section come with a degree of generality and are not restricted to only the previously defined Markov chain for the case  $n = 4$ , but will also be used in an identical manner for an alternatively defined Markov chain with  $n > 4$ . Lemma 9 is independent of any previous definitions and holds for any adapted processes  $K_{i,t}$  satisfying the inequality (2.6), regardless of our definition of  $K_{i,t}$  as a function of  $(u^t, v^t)$ . Lemma 10 is a general observation, independent of anything discussed so far. Lemma 11 relies on the same premise as Lemma 9, but here we do make use of the fact that  $K_{i,t}$  is a function of  $(u^t, v^t)$ . The definition and dimension of the random variable  $(u^t, v^t)$  is irrelevant in the proof, and we only require that it satisfies the Markov property. Lemma 12 makes explicit reference to the process  $D_t$  defined in (2.8), however the result remains true for any similar process used to bounding  $R_t$  in a similar manner, for any  $R_t$  satisfying the necessary preconditions. Lastly, the transition from Lemma 12 to Corollary 13 is only based on some simplification of binomial coefficients, hence it will also translate easily to the case  $n > 4$ .

To illustrate the relevance of these results to the problem we have built so far (i.e. for  $n = 4$ ), we will start by assuming the existence of a set of auxiliary processes which satisfy conditions outlined below, and which can serve to provide an upper bound to the random part of  $\dot{r}_t$ , namely to  $\max \left\{ \left(\frac{4\mu_1}{\mu_2} + 4\right) \left(\frac{u_2^t}{x} + \frac{x}{v_2^t} + 2\right), 4 + \frac{4\mu_1}{bv_4^t} \right\}$ .

Suppose that for  $i = 1, \dots, N$ , the processes  $K_{i,t} = K(u^t, v^t)$  are adapted to  $\mathcal{F}_t$ , and that for  $t \geq 0$

$$(2.6) \quad \mathbb{E}[K_{i,t+1} | \mathcal{F}_t] \leq \zeta_i K_{i,t} + c_i$$

where  $\zeta_i < 1$  and  $c_i$  are constants. Then for the process  $J_t = J(u^t, v^t) := \sum_i K_{i,t}$  we also have

$$\mathbb{E}[J_{t+1} | \mathcal{F}_s] \leq \max \{\zeta_i\} J_t + \sum_i c_i$$

In particular, if  $J_s \geq \eta := \frac{2 \sum c_i}{1 - \max \{\zeta_i\}}$  for some  $s \geq 0$ , then  $\mathbb{E}[J_{s+1} | \mathcal{F}_s] \leq \beta J_s$  where  $\beta := \frac{1 + \max \{\zeta_i\}}{2}$ . Let

$T = T(s) := \min \{ \tau > s \mid J_\tau \leq \eta \}$ , and for  $s \geq 0$  and  $t \geq 1$  define

$$\hat{J}_{s,t} := \begin{cases} J_{s+t} & s+t < T, J_s \leq \eta \\ 0 & \text{otherwise} \end{cases}$$

or in other words  $\hat{J}_{s,t} = \mathbf{1}_{\{J_s \leq \eta\} \cap \{T > s+t\}} J_{s+t}$ .

**Lemma 9.** *For the notation and assumptions of the preceding paragraph,  $\mathbb{E}[\hat{J}_{s,t+1} \mid \mathcal{F}_s] \leq \beta^{t+1} \eta$  for  $t \geq 0$  and  $s \geq 0$ .*

*Proof.* Observe that for  $t \geq 1$ ,

$$\begin{aligned} \mathbb{E}[\hat{J}_{s,t+1} \mid \mathcal{F}_{s+t}] &= \mathbf{1}_{\{J_s \leq \eta\} \cap \{T \leq s+t\}} \mathbb{E}[\hat{J}_{s,t+1} \mid \mathcal{F}_{s+t}] + \mathbf{1}_{\{J_s \leq \eta\} \cap \{T > s+t\}} \mathbb{E}[\hat{J}_{s,t+1} \mid \mathcal{F}_{s+t}] \\ &= 0 + \mathbf{1}_{\{J_s \leq \eta\} \cap \{T > s+t\}} \mathbb{E}[\mathbf{1}_{\{T > s+t+1\}} J_{s+t+1} \mid \mathcal{F}_{s+t}] \\ &\leq \mathbf{1}_{\{J_s \leq \eta\} \cap \{T > s+t\}} \mathbb{E}[\mathbf{1}_{\{J_{s+t} > \eta\}} J_{s+t+1} \mid \mathcal{F}_{s+t}] \\ &\leq \mathbf{1}_{\{J_s \leq \eta\} \cap \{T > s+t\}} \beta J_{s+t} \\ &= \beta \hat{J}_{s,t} \end{aligned}$$

Proceeding inductively, it follows that

$$\mathbb{E}[\hat{J}_{s,t+1} \mid \mathcal{F}_s] \leq \mathbb{E}[\beta^t \hat{J}_{s,1} \mid \mathcal{F}_s]$$

Finally,

$$\begin{aligned} \mathbb{E}[\hat{J}_{s,1} \mid \mathcal{F}_s] &\leq \mathbb{E}[\mathbf{1}_{\{J_s \leq \eta\}} J_{s+1} \mid \mathcal{F}_s] \\ &= \mathbb{E}[J_{s+1} \mid \mathcal{F}_s] \mathbf{1}_{\{J_s \leq \eta\}} \\ &\leq \left( \sum_i c_i + \max \{ \zeta_i \} J_s \right) \mathbf{1}_{\{J_s \leq \eta\}} \\ &\leq \left( \sum_i c_i + \max \{ \zeta_i \} \eta \right) \mathbf{1}_{\{J_s \leq \eta\}} \\ &\leq \beta \eta \end{aligned}$$

□

*Remark.* If it is uncertain that  $J_s \leq \eta$ , we can still define  $\check{J}_{s,t} = \mathbf{1}_{\{T > s+t\}} J_{s+t}$ , and following the proof of Lemma 9 it is a straight forward conclusion that

$$(2.7) \quad \mathbb{E}[\check{J}_{s,t+1} \mid \mathcal{F}_s] \leq \beta^{t+1} \max \{ \eta, J_s \}$$

Now suppose that  $D_t$  is a process adapted to  $\mathcal{F}_t$  such that for  $t \geq 1$ ,  $D_t \geq \max \left\{ \left( \frac{4\mu_1}{\mu_2} + 4 \right) \left( \frac{u_2^t}{x} + \frac{x}{v_2^t} + 2 \right), 4 + \frac{4\mu_1}{bv_4^t} \right\}$ . Furthermore, suppose also that

$$(2.8) \quad D_{t+1} \leq \omega_{N+1,t+1} + \sum_{i=1}^N \omega_{i,t+1} K_{i,t}$$

where  $(\omega_{1,t+1}, \dots, \omega_{N+1,t+1})$  is a non-negative random vector, i.i.d. over time  $t \geq 1$ , measurable w.r.t.  $\mathcal{F}_{t+1}$  and independent of  $\mathcal{F}_t$ . It is now clear that  $D_t$  is defined with the intent to serve as an upper bound for  $\hat{r}_t$ . We will construct  $D_t$  in Section 2.1.2, and the reasons for insisting on the condition given in (2.8) will become apparent.

If  $S$  is a finite a.s. stopping time adapted to  $\mathcal{F}_t$  s.t.  $J_S \leq \eta$ , then  $D_{S+1} \leq \eta \sum \omega_{i,S+1} + \omega_{N+1,S+1}$ . Therefore, applying Lemma 8 we get

$$\begin{aligned}
\mathbb{E}[R_{S+2}] &\leq \mathbb{E}[Q_{S+1}R_{S+1}] \\
&\leq \mathbb{E}[\dot{r}_{S+1}(R_{S+1} - 1)] + 1 \\
&\leq \mathbb{E}\left[\left(1 - \frac{1}{D_{S+1}}\right)(R_{S+1} - 1)\right] + 1 \\
&\leq \mathbb{E}\left[\left(1 - \frac{1}{D_{S+1}}\right)(R_S - 1)\right] + 1 \\
&\leq \mathbb{E}\left[\left(1 - \frac{1}{\eta \sum \omega_{i,S+1} + \omega_{N+1,S+1}}\right)(R_S - 1)\right] + 1 \\
&= \mathbb{E}\left[\left(1 - \frac{1}{\eta \sum \omega_{i,S+1} + \omega_{N+1,S+1}}\right)\right] \mathbb{E}[(R_S - 1)] + 1 \\
(2.9) \quad &\leq r \mathbb{E}[(R_S - 1)] + 1
\end{aligned}$$

where  $r = 1 - 1/((\theta_1 + \dots + \theta_N)\eta + \theta_{N+1})$  and  $\theta_i := \mathbb{E}[\omega_{i,t+1}]$ . Here we have used Jensen's inequality in the transition between the last two lines in (2.9). An additional observation that will be useful to us later, is that if  $0 \leq Y \in \mathcal{F}_s$  then by a derivation identical to (2.9) we get

$$(2.10) \quad \mathbb{E}[Y R_{S+2}] \leq r \mathbb{E}[Y (R_S - 1)] + \mathbb{E}[Y]$$

The term  $\mathbb{E}[Y]$  in the right-hand side of (2.10) comes about in the second line of (2.9), as a result of applying (2.5).

**Lemma 10.** *Let  $Y$  be a random variable. If  $A$  is an event and  $B \subseteq \mathbb{R}$ ,*

$$\mathbb{P}[A | Y \in B] \leq \sup_{y_0 \in B} \mathbb{P}[A | Y = y_0]$$

*Proof.* Suppose on the contrary that  $\mathbb{P}[A | Y \in B] > \mathbb{P}[A | Y = y_0]$  for all  $y_0$  in  $B$ . Multiplying both sides by the marginal density  $f_Y(y_0)$  and integrating over  $y_0$  in  $B$  gives the contradiction

$$\mathbb{P}[A, Y \in B] > \mathbb{P}[A, Y \in B]$$

□

Working under the assumption that a process  $D_t$  satisfying the aforementioned conditions exists, and that Lemma 9 applies, we define the set  $\bar{S}_t$  by

$$(2.11) \quad \bar{S}_t := \{1 \leq i \leq t | J_i \leq \eta\}$$

Then

**Lemma 11.** *For any subset  $\{c_1, c_2, \dots, c_k\} \subseteq \{1, \dots, t\}$ ,  $\mathbb{P}[\bar{S}_t = \{c_1, c_2, \dots, c_k\} | J_0 \leq \eta] \leq \beta^{t-k}$*

For the purpose of this lemma we will consider the previously defined function  $J$  as a function on  $\mathbb{R}_+^8$ , as this will allow us to refer to the odd numbered co-ordinates  $u_1^t, u_3^t, v_1^t$  and  $v_3^t$  when we later define a satisfactory auxiliary process. It will be clear that this interpretation has no impact on any of previously derived results (such as Lemma 9) that we may need to refer to.

*Proof.* Let  $A = \{(y_1, \dots, y_8) \in \mathbb{R}_+^8 \text{ s.t. } J(y_1, \dots, y_8) \leq \eta\}$ , and  $I \subseteq \{0, 1, \dots, k\}$  be those indices  $i$  that satisfy  $c_{i+1} > c_i + 1$ , where by convention we set  $c_0 = 0$  and  $c_{k+1} = t + 1$ . Then for  $i \in I$  let  $B_i = \{J_{c_{i+1}} > \eta, \dots, J_{c_{i+1}-1} > \eta\}$ . By Lemma 10,

$$\mathbb{P}[B_i | J_{c_i} \leq \eta] \leq \sup_{y \in A} \mathbb{P}[B_i | (u^{c_i}, v^{c_i}) = y]$$

and since  $J_{c_i}$  is determined by the values  $(u^{c_i}, v^{c_i})$ , it follows by the same reasoning and the Markov property that also for any event  $C_{c_i-1} \in \mathcal{F}_{c_i-1}$

$$(2.12) \quad \mathbb{P}[B_i | J_{c_i} \leq \eta, C_{c_i-1}] \leq \sup_{y \in A} \mathbb{P}[B_i | (u^{c_i}, v^{c_i}) = y]$$

Observe also that if  $I = \{i_1, \dots, i_m\}$  for some  $m \leq k+1$ , then

$$\begin{aligned} \sum_{j=1}^m (c_{i_j+1} - c_{i_j} - 1) &= |\{1, \dots, t\} \setminus \{c_1, c_2, \dots, c_k\}| \\ &= t - k \end{aligned}$$

Hence we get

$$\begin{aligned} \mathbb{P}[\bar{S}_t = \{c_1, c_2, \dots, c_k\} | J_0 \leq \eta] &= \mathbb{P}[\{J_{c_1} \leq \eta, \dots, J_{c_k} \leq \eta\} \cap \{\cap_{i \in I} B_i\} | J_0 \leq \eta] \\ &\leq \mathbb{P}[\{J_{c_{i_1}} \leq \eta, \dots, J_{c_{i_m}} \leq \eta\} \cap \{\cap_{i \in I} B_i\} | J_0 \leq \eta] \\ &= \mathbb{P}[B_{i_m} | \{J_{c_{i_1}} \leq \eta, \dots, J_{c_{i_m}} \leq \eta\} \cap \{\cap_{j=1}^{m-1} B_{i_j}\}, J_0 \leq \eta] \cdot \\ &\quad \mathbb{P}[\{J_{c_{i_1}} \leq \eta, \dots, J_{c_{i_m}} \leq \eta\} \cap \{\cap_{j=1}^{m-1} B_{i_j}\} | J_0 \leq \eta] \\ &\leq \sup_{y \in A} \mathbb{P}[B_{i_m} | (u^{c_{i_m}}, v^{c_{i_m}}) = y] \cdot \mathbb{P}[\{J_{c_{i_1}} \leq \eta, \dots, J_{c_{i_{m-1}}} \leq \eta\} \cap \{\cap_{j=1}^{m-1} B_{i_j}\} | J_0 \leq \eta] \\ &\quad \vdots \\ &\leq \prod_{j=1}^m \sup_{y \in A} \mathbb{P}[B_{i_j} | (u^{c_{i_j}}, v^{c_{i_j}}) = y] \\ &\leq \prod_{j=1}^m \sup_{y \in A} \mathbb{P}[\hat{J}_{c_{i_j}, c_{(i_j+1)} - c_{i_j} - 1} > \eta | (u^{c_{i_j}}, v^{c_{i_j}}) = y] \\ &\leq \beta^{c_{i_1+1} - c_{i_1} - 1} \dots \beta^{c_{i_m+1} - c_{i_m} - 1} \\ &= \beta^{t-k} \end{aligned}$$

The inequality before the last line follows from Lemma 9 and Markov's inequality. We note that when  $i_1 = 0$ , the event  $J_{i_1} \leq \eta$  appears in the second line, but not in the first. This can be justified by observing that in this case  $J_{i_1} = J_0$ , and  $J_0 \leq \eta$  is already given in this conditional probability.  $\square$

From Lemma 11 it also immediately follows that

$$(2.13) \quad \mathbb{P}[|\bar{S}_t| = k | J_0 \leq \eta] \leq \binom{t}{k} \beta^{t-k}$$

Assuming now that the process  $D_t$  defined in (2.8) exists, we can conclude the following in the event  $\{J_0 \leq \eta\}$ :

**Lemma 12.**  $\mathbb{E} \left[ R_{t+2} \mathbf{1}_{|\bar{S}_t| > k} | \mathcal{F}_0 \right] - \mathbb{P} [|\bar{S}_t| > k | \mathcal{F}_0] \leq r^{\lceil (k+1)/2 \rceil} (R_0 - 1)$

*Proof.* Let  $\tau_0 = 0$  and  $\{\tau_i\} \subseteq \{1, 2, \dots\}$  be those times for which  $J_{\tau_i} \leq \eta$ . Then by (2.10) with  $Y = \mathbf{1}_{\tau_{k+1} \leq t}$  and  $S = \tau_{k+1}$

$$\begin{aligned} \mathbb{E} \left[ R_{t+2} \mathbf{1}_{|\bar{S}_t| > k} \mid \mathcal{F}_0 \right] &= \mathbb{E} \left[ R_{t+2} \mathbf{1}_{\tau_{k+1} \leq t} \mid \mathcal{F}_0 \right] \\ &\leq \mathbb{E} \left[ R_{\tau_{k+1}+2} \mathbf{1}_{\tau_{k+1} \leq t} \mid \mathcal{F}_0 \right] \\ &\leq r \mathbb{E} \left[ \mathbf{1}_{\tau_{k+1} \leq t} (R_{\tau_{k+1}} - 1) \mid \mathcal{F}_0 \right] + \mathbb{P} \left[ |\bar{S}_t| > k \mid \mathcal{F}_0 \right] \\ &\leq r \mathbb{E} \left[ \mathbf{1}_{\tau_{k-1} \leq t} (R_{\tau_{k-1}+2} - 1) \mid \mathcal{F}_0 \right] + \mathbb{P} \left[ |\bar{S}_t| > k \mid \mathcal{F}_0 \right] \end{aligned}$$

The last inequality uses the fact that  $\mathbf{1}_{\tau_{k+1} \leq t} \leq \mathbf{1}_{\tau_{k-1} \leq t}$  and  $R_{\tau_{k+1}} \leq R_{\tau_{k-1}+2}$ . This then leads to the first step in an inductive argument:

$$(2.14) \quad \mathbb{E} \left[ R_{\tau_{k+1}+2} \mathbf{1}_{\tau_{k+1} \leq t} \mid \mathcal{F}_0 \right] - \mathbb{P} \left[ |\bar{S}_t| > k \mid \mathcal{F}_0 \right] \leq r \left( \mathbb{E} \left[ R_{\tau_{k-1}+2} \mathbf{1}_{\tau_{k-1} \leq t} \mid \mathcal{F}_0 \right] - \mathbb{P} \left[ |\bar{S}_t| > k-2 \mid \mathcal{F}_0 \right] \right)$$

Proceeding in this manner, we claim that we get

$$\mathbb{E} \left[ R_{\tau_{k+1}+2} \mathbf{1}_{\tau_{k+1} \leq t} \mid \mathcal{F}_0 \right] - \mathbb{P} \left[ |\bar{S}_t| > k \mid \mathcal{F}_0 \right] \leq r^{\lceil (k+1)/2 \rceil} (R_0 - 1)$$

The ceiling function in the exponent  $\lceil (k+1)/2 \rceil$  is immediate whenever  $k+1$  is even. If on the other hand  $k+1$  is odd, by (2.14)

$$\begin{aligned} \mathbb{E} \left[ R_{\tau_{k+1}+2} \mathbf{1}_{\tau_{k+1} \leq t} \mid \mathcal{F}_0 \right] - \mathbb{P} \left[ |\bar{S}_t| > k \mid \mathcal{F}_0 \right] &\leq r^{\lfloor (k+1)/2 \rfloor} \mathbb{E} \left[ \mathbf{1}_{\tau_1 \leq t} (R_{\tau_1+2} - 1) \mid \mathcal{F}_0 \right] \\ &\leq r^{\lfloor (k+1)/2 \rfloor} r \mathbb{E} \left[ \mathbf{1}_{\tau_1 \leq t} (R_{\tau_1} - 1) \mid \mathcal{F}_0 \right] \\ &\leq r^{\lfloor (k+1)/2 \rfloor + 1} (R_0 - 1) \end{aligned}$$

The second line follows from (2.10). □

From (2.13) and Lemma 12 we conclude

$$\begin{aligned} \mathbb{E} [R_{t+2} \mid \mathcal{F}_0] &= \mathbb{E} \left[ R_{t+2} \mathbf{1}_{|\bar{S}_t| > k} \mid \mathcal{F}_0 \right] + \mathbb{E} \left[ R_{t+2} \mathbf{1}_{|\bar{S}_t| \leq k} \mid \mathcal{F}_0 \right] \\ &\leq r^{\lceil (k+1)/2 \rceil} (R_0 - 1) + \mathbb{P} \left[ |\bar{S}_t| > k \mid \mathcal{F}_0 \right] + \mathbb{E} \left[ R_0 \mathbf{1}_{|\bar{S}_t| \leq k} \mid \mathcal{F}_0 \right] \\ &\leq r^{\lceil (k+1)/2 \rceil} (R_0 - 1) + \mathbb{P} \left[ |\bar{S}_t| > k \mid \mathcal{F}_0 \right] + (R_0 - 1) \mathbb{P} \left[ |\bar{S}_t| \leq k \mid \mathcal{F}_0 \right] + \mathbb{P} \left[ |\bar{S}_t| \leq k \mid \mathcal{F}_0 \right] \\ (2.15) \quad &\leq 1 + (R_0 - 1) \left( r^{\lceil (k+1)/2 \rceil} + \sum_{j=0}^k \binom{t}{j} \beta^{t-j} \right) \end{aligned}$$

Inequality (2.15) is true for any  $k \leq t$ , so we are free to choose any value for  $k$  in this range. We can simplify this expression by removing the binomial terms in the following manner: note first that if we take  $k \leq \lfloor \frac{t}{3} \rfloor$  then for  $j < k$ ,  $\binom{t}{j} \leq \frac{1}{2} \binom{t}{j+1}$ , hence  $\sum_{j=0}^k \binom{t}{j} \beta^{t-j} \leq 2 \binom{t}{k} \beta^{t-k}$ . Next, let  $d := \lceil \frac{t}{k} \rceil$  and note that  $\binom{dk}{k} q^k (1-q)^{dk-k} \leq 1$  whenever  $q \in (0, 1)$ . Therefore, if  $d \geq 2$  then by taking  $q = \frac{1}{d}$  we get

$$(2.16) \quad \binom{t}{k} \leq \binom{dk}{k} \leq \frac{d^{dk}}{(d-1)^{(d-1)k}}$$

From these remarks and conditions, it follows that the summation in (2.15) may be replaced by  $2 \left( \frac{d^d}{(d-1)^{(d-1)}} \right)^k \beta^{t-k}$ . Our goal is to bound  $d$  from below by a constant  $d'$  (and thereby set  $k$  to be a fraction of  $t$ ) in such a way that  $\left( \frac{d^d}{(d-1)^{(d-1)}} \right)^k \beta^{t-k}$  is decaying exponentially in  $t$ . Since  $\left( \frac{d^d}{(d-1)^{(d-1)}} \right)^k \beta^{t-k} = \left( \left( \frac{d^d}{(d-1)^{(d-1)}} \right) \beta^{t/k-1} \right)^k$ , this aim would be achieved if we could find  $d'$  such that for  $d \geq d'$

$$(2.17) \quad \frac{d^d}{(d-1)^{(d-1)}} \beta^{(d-2)} \leq \sqrt{r}$$

Here the term  $\sqrt{r}$  is chosen for convenience, and (2.17) would then imply that

$$\sum_{j=0}^k \binom{t}{j} \beta^{t-j} \leq 2 \left( \left( \frac{d^d}{(d-1)^{(d-1)}} \right) \beta^{t/k-1} \right)^k \leq 2r^{k/2}$$

The left hand side of (2.17) is equal to  $d \left( 1 + \frac{1}{d-1} \right)^{d-1} \beta^{(d-2)} \leq ed\beta^{(d-2)}$ . Hence (2.17) is true if  $d \geq \frac{\ln(\frac{\sqrt{r}}{ed})}{\ln(\beta)} + 2 = \frac{1}{|\ln(\beta)|} \left\{ \ln(d) + \ln\left(\frac{e}{\sqrt{r}}\right) \right\} +$

2. Since  $\ln(d) \leq \sqrt{d}$ , we can consider the inequality  $d \geq \frac{1}{|\ln(\beta)|} \sqrt{d} + \frac{\ln(\frac{e}{\sqrt{r}})}{|\ln(\beta)|} + 2$  or  $\sqrt{d} \geq \frac{1}{|\ln(\beta)|} + \frac{\ln(\frac{e}{\sqrt{r}})}{|\ln(\beta)|\sqrt{d}} + \frac{2}{\sqrt{d}}$ , which is certainly true if  $d \geq \left( \frac{1}{|\ln(\beta)|} + \frac{1}{|\ln(\beta)|} \ln\left(\frac{e}{\sqrt{r}}\right) + 2 \right)^2$ . Therefore taking  $d' := \left( \frac{1}{|\ln(\beta)|} + \frac{1}{|\ln(\beta)|} \ln\left(\frac{e}{\sqrt{r}}\right) + 2 \right)^2$  and setting  $k = \lfloor \frac{t}{d'} \rfloor$  (note that the condition  $k \leq \lfloor \frac{t}{3} \rfloor$  from the previous paragraph is satisfied), we get that  $d = \lceil \frac{t}{k} \rceil \geq d'$ , hence

$$\begin{aligned} 1 + (R_0 - 1) \left( r^{\lceil (k+1)/2 \rceil} + \sum_{j=0}^k \binom{t}{j} \beta^{t-j} \right) &\leq 1 + (R_0 - 1) \left( r^{\lceil (k+1)/2 \rceil} + 2r^{k/2} \right) \\ (2.18) \qquad \qquad \qquad &\leq 1 + r^{\frac{1}{2} \lfloor \frac{t}{d'} \rfloor} (R_0 - 1) (\sqrt{r} + 2) \end{aligned}$$

We summarise this in the following corollary

**Corollary 13.**  $\mathbb{E}[R_{t+2} | \mathcal{F}_0] \leq 1 + 3r^{\frac{1}{2} \lfloor \frac{t}{d'} \rfloor} (R_0 - 1)$  for  $t \geq d'$ , where  $d'$  and  $r$  are as defined previously.

2.1.2. *Construction of  $D_t$ .* For ease of reference, we will start by first giving the following list of definitions

$$\begin{aligned} K_{1,t} &:= u_2^t + u_4^t & K_{2,t} &:= \frac{u_3^t + u_1^t + b}{\gamma_2^t + \gamma_4^t} \\ D_t &:= \frac{1}{x} \left( \frac{4\mu_1}{\mu_2} + 4 \right) (u_2^t + u_4^t) + \left( \left( \frac{4\mu_1}{\mu_2} + 4 \right) x + \frac{4\mu_1}{b} \right) \left( \frac{1}{u_2^t} + \frac{1}{u_4^t} \right) & \zeta_2 &:= \frac{a_3 + a_4}{a_2 + a_3 + a_4 + a_5 - 1} \\ \zeta_1 &:= \frac{a_2 + a_3}{a_1 + a_2 + a_3 + a_4 - 1} & C_2 &:= \frac{a_1 + a_2 + x b}{x(a_2 + a_3 + a_4 + a_5 - 1)} \\ C_1 &:= \zeta_1 x + \frac{a_4 + a_5}{b} & \omega_{1,t+1} &:= \frac{1}{x} \left( \frac{4\mu_1}{\mu_2} + 4 \right) \frac{\gamma_2^{t+1}}{\gamma_1^{t+1} + \gamma_3^{t+1}} \\ \tilde{\omega}_{2,t+1} &= 2 + \frac{\gamma_2^{t+1}}{\gamma_4^{t+1}} + \frac{\gamma_4^{t+1}}{\gamma_2^{t+1}} & \omega_{2,t+1} &:= \left( \left( \frac{4\mu_1}{\mu_2} + 4 \right) x + \frac{4\mu_1}{b} \right) \tilde{\omega}_{2,t+1} \frac{\gamma_3^{t+1}}{\gamma_2^{t+1} + \gamma_4^{t+1}} \\ \omega_{2,t+1} &:= \left( \left( \frac{4\mu_1}{\mu_2} + 4 \right) x + \frac{4\mu_1}{b} \right) \tilde{\omega}_{2,t+1} \frac{\gamma_3^{t+1}}{\gamma_2^{t+1} + \gamma_4^{t+1}} & \omega_{3,t+1} &:= \frac{1}{x} \left( \frac{4\mu_1}{\mu_2} + 4 \right) \left( \frac{\gamma_2^{t+1}}{\gamma_1^{t+1} + \gamma_3^{t+1}} x + \frac{\gamma_4^{t+1}}{b} \right) + \left( \left( \frac{4\mu_1}{\mu_2} + 4 \right) x + \frac{4\mu_1}{b} \right) \tilde{\omega}_{2,t+1} \frac{\gamma_1^{t+1} + b}{\gamma_2^{t+1} + \gamma_4^{t+1}} \end{aligned}$$

Note that

$$(2.19) \qquad \max \left\{ \left( \frac{4\mu_1}{\mu_2} + 4 \right) \left( \frac{u_2^t}{x} + \frac{x}{v_2^t} + 2 \right), 4 + \frac{4\mu_1}{b v_4^t} \right\} \leq D_t$$

where we have used the fact that  $u \preceq v$  and  $2 \leq \frac{u}{x} + \frac{x}{u}$ . To bound the first term in this sum, observe that

$$\begin{aligned} u_2^{t+1} + u_4^{t+1} &= \frac{\gamma_2^{t+1}}{\frac{\gamma_1^{t+1}}{x + u_2^t} + \frac{\gamma_3^{t+1}}{u_2^t + u_4^t}} + \frac{\gamma_4^{t+1}}{\frac{\gamma_3^{t+1}}{u_2^t + u_4^t} + b} \\ (2.20) \qquad \qquad \qquad &\leq \frac{\gamma_2^{t+1}}{\gamma_1^{t+1} + \gamma_3^{t+1}} (u_2^t + u_4^t + x) + \frac{\gamma_4^{t+1}}{b} \end{aligned}$$

Therefore  $\mathbb{E}[K_{1,t+1} | \mathcal{F}_t] \leq \zeta_1 K_{1,t} + C_1$ . Observe that since

$$\begin{aligned} u_3^{t+1} &= \frac{\gamma_3^{t+1}}{u_2^t + u_4^t} = \frac{\gamma_3^{t+1}}{\frac{\gamma_2^t}{u_1^t + u_3^t} + \frac{\gamma_4^t}{u_3^t + b}} \\ &\leq \frac{\gamma_3^{t+1}}{\gamma_2^t + \gamma_4^t} (u_1^t + u_3^t + b) \end{aligned}$$

it follows that

$$(2.21) \quad K_{2,t+1} \leq \frac{\gamma_3^{t+1}}{\gamma_2^{t+1} + \gamma_4^{t+1}} K_{2,t} + \frac{u_1^{t+1} + b}{\gamma_2^{t+1} + \gamma_4^{t+1}}$$

and hence

$$\begin{aligned} (2.22) \quad \mathbb{E}[K_{2,t+1} | \mathcal{F}_t] &\leq \zeta_2 K_{2,t} + \mathbb{E}\left[\frac{\frac{\gamma_1^{t+1}}{x} + b}{\gamma_2^{t+1} + \gamma_4^{t+1}} | \mathcal{F}_t\right] \\ &\leq \zeta_2 K_{2,t} + C_2 \end{aligned}$$

Both  $K_{1,t}$  and  $K_{2,t}$  are adapted to  $\mathcal{F}_t$  and are in fact functions of  $u^t$  (since  $\gamma_2^t + \gamma_4^t = u_2^t (u_1^t + u_3^t) + u_4^t (u_3^t + b)$ ). Note also that

$$\begin{aligned} (2.23) \quad \frac{1}{u_2^{t+1}} + \frac{1}{u_4^{t+1}} &\leq \left( \frac{1}{\gamma_2^{t+1}} + \frac{1}{\gamma_4^{t+1}} \right) (u_1^{t+1} + u_3^{t+1} + b) \\ &= \tilde{\omega}_{2,t+1} K_{2,t+1} \end{aligned}$$

and  $\tilde{\omega}_{2,t+1}$  is independent of  $\mathcal{F}_t$ . By (2.20), (2.21) and (2.23) we conclude that

$$\begin{aligned} D_{t+1} &\leq \frac{1}{x} \left( \frac{4\mu_1}{\mu_2} + 4 \right) \left( \frac{\gamma_2^{t+1}}{\gamma_1^{t+1} + \gamma_3^{t+1}} (K_{1,t} + x) + \frac{\gamma_4^{t+1}}{b} \right) + \left( \left( \frac{4\mu_1}{\mu_2} + 4 \right) x + \frac{4\mu_1}{b} \right) \tilde{\omega}_{2,t+1} \left( \frac{\gamma_3^{t+1}}{\gamma_2^{t+1} + \gamma_4^{t+1}} K_{2,t} + \frac{\frac{\gamma_1^{t+1}}{x} + b}{\gamma_2^{t+1} + \gamma_4^{t+1}} \right) \\ &\leq \omega_{1,t+1} K_{1,t} + \omega_{2,t+1} K_{2,t} + \omega_{3,t+1} \end{aligned}$$

and hence  $D_t$  satisfies the conditions given by and preceding equation (2.8). Referring back to (2.9), we obtain the rate

$$(2.24) \quad r = 1 - \frac{1}{(\theta_1 + \theta_2)\eta + \theta_3}$$

where  $\theta_1, \theta_2, \theta_3$  are the expected values of  $\omega_{1,t+1}, \omega_{2,t+1}$  and  $\omega_{3,t+1}$  respectively.

We make the additional note that it is not necessary for  $\{K_{i,t}\}$  to be deterministic functions of  $(u^t, v^t)$ . This assumption was required to make use of the Markov property in (2.9) and (2.12), however the arguments remain true if  $\{K_{i,t}\}$  are random functions of  $(u^t, v^t)$  with random terms that are independent of  $\mathcal{F}_\infty$ .

Note also that condition (1.6) guarantees that  $\zeta_1 < 1$  and  $\zeta_2 < 1$ , as well as the finite value of all constants and finite expectation of all random variables defined in the beginning of this section.

We have now established a sufficient foundation to prove our first theorem.

*Proof of Theorem 1.* By Corollary 5,  $\mathbb{P}[u^{t+3} \neq v^{t+3}]$  is an upper bound for  $d_{TV}(\mathcal{U}^{t+3}, \mathcal{V}^{t+3})$  under the specified 'one shot' coupling described in the paragraph preceding the corollary (in the aforementioned description we couple uniformly until time  $t$ , and attempt to merge the two Markov chains thereafter. Here we attempt to do this after time  $t+2$ , but the argument remains the same). In the event  $\{J_0 \leq \eta\}$  we conclude by Corollary 13

$$\mathbb{E}[R_{t+2} - 1 | \mathcal{F}_0, ] \leq 3r^{\frac{1}{2}} \lfloor \frac{t}{d} \rfloor (R_0 - 1)$$

Therefore by Lemma 6 and using Jensen's inequality

$$\begin{aligned}
 \mathbb{P}[u^{t+3} \neq v^{t+3} | \mathcal{F}_0] &= \mathbb{E}[\mathbb{P}[u^{t+3} \neq v^{t+3} | \mathcal{F}_{t+2}] | \mathcal{F}_0,] \\
 &\leq \mathbb{E}\left[1 - (R_{t+2})^{-(a_2+a_3+a_4+a_5)}\right] \\
 &\leq 1 - (\mathbb{E}[R_{t+2}])^{-(a_2+a_3+a_4+a_5)} \\
 &\leq 1 - \left(1 + 3r^{\frac{1}{2}\lfloor \frac{t}{2d'} \rfloor} (R_0 - 1)\right)^{-(a_2+a_3+a_4+a_5)}
 \end{aligned}
 \tag{2.25}$$

We claim that the right-hand side of (2.25) is bounded by  $3r^{\frac{1}{2}\lfloor \frac{t}{2d'} \rfloor} (a_2 + a_3 + a_4 + a_5) (R_0 - 1)$ . To justify this claim, define  $E(y) := \frac{1}{(1+y)^\nu} + \nu y$  for  $y, \nu \in \mathbb{R}^+$ , and observe that  $E'(y) = \frac{-\nu}{(1+y)^{\nu+1}} + \nu \geq 0$ . Hence  $E(y) \geq E(0) = 1$ . Now take  $\nu = a_2 + a_3 + a_4 + a_5$  and  $y = 3r^{\frac{1}{2}\lfloor \frac{t}{2d'} \rfloor} (R_0 - 1)$ , and we get

$$\frac{1}{\left(1 + 3r^{\frac{1}{2}\lfloor \frac{t}{2d'} \rfloor} (R_0 - 1)\right)^{a_2+a_3+a_4+a_5}} + 3(a_2 + a_3 + a_4 + a_5) r^{\frac{1}{2}\lfloor \frac{t}{2d'} \rfloor} (R_0 - 1) \geq 1$$

The claim now follows immediately, as does the first statement of the theorem. If we are no longer restricted to the event  $\{J_0 \leq \eta\}$ , then (recall that  $T$  is the first time  $t$  such that  $J_t \leq \eta$ ) by the remark (2.7)

$$\begin{aligned}
 \mathbb{P}[u^{t+3} \neq v^{t+3} | \mathcal{F}_0] &\leq \mathbb{P}\left[u^{t+3} \neq v^{t+3} \mid J_0 > \eta, T \leq \left\lfloor \frac{t}{2} \right\rfloor + 3\right] + \mathbb{P}\left[T > \left\lfloor \frac{t}{2} \right\rfloor + 3 \mid J_0 > \eta\right] \\
 &\leq r^{\frac{1}{2}\lfloor \frac{t}{2d'} \rfloor} (1 + 3(a_2 + a_3 + a_4 + a_5) (R_0 - 1)) + \frac{\max\{J_0, \eta\} \beta^{\lfloor \frac{t}{2} \rfloor + 3}}{\eta}
 \end{aligned}
 \tag{2.26}$$

Since this is greater than what we have on  $\{J_0 \leq \eta\}$ , it is also a bound for general values of  $J_0$ .  $\square$

**2.1.3. Sampling from equilibrium.** One application of the previously derived results lies in sampling from the equilibrium distribution  $\pi$  defined by (1.2). We will start by taking  $\mathcal{U}^0 = (1, 1, 1, 1)$  and  $\mathcal{V}^0 \sim \pi$ , and define  $u_0, v_0$  according to (1.7). Then from Theorem 1, using the one-shot coupling at time  $t + 3$ , it follows that for  $t \geq d'$

$$\mathbb{P}[u^{t+3} \neq v^{t+3} | \mathcal{F}_0] \leq r^{\frac{1}{2}\lfloor \frac{t}{2d'} \rfloor} (1 + 3(a_2 + a_3 + a_4 + a_5) (R_0 - 1)) + \frac{\max\{J_0, \eta\} \beta^{\lfloor \frac{t}{2} \rfloor + 3}}{\eta}$$

*Proof of Corollary 2.* This follows immediately from Theorem 1.  $\square$

Now let  $C_g := \int \left(\prod_{i=1}^4 z_i^{a_i+a_{i+1}-1}\right) \exp\left(\sum_{i=1}^5 -z_i z_{i-1}\right) dz$ . Then we can bound the terms  $\mathbb{E}_\pi[R_0]$  and  $\mathbb{E}_\pi[J_0]$  in Corollary 2 in the following way:

$$\begin{aligned}
 d_{TV}(\mathcal{U}^{t+3}, \pi) &\leq \mathbb{P}[u^{t+3} \neq v^{t+3}] \\
 &\leq r^{\frac{1}{2}\lfloor \frac{t}{2d'} \rfloor} \frac{1}{C_g} (1 + 3(a_2 + a_3 + a_4 + a_5)) \int \left(\frac{\max_i\{1, v_i\}}{\min_i\{1, v_i\}}\right) \left(\prod_{i=1}^4 v_i^{a_i+a_{i+1}-1}\right) \exp\left(\sum_{i=1}^5 -v_i v_{i-1}\right) dv \\
 &\quad + \frac{1}{\eta} \beta^{\lfloor \frac{t}{2} \rfloor + 3} \left(\eta + \frac{1}{C_g} \int J_0 \left(\prod_{i=1}^4 v_i^{a_i+a_{i+1}-1}\right) \exp\left(\sum_{i=1}^5 -v_i v_{i-1}\right) dv\right) \\
 &\leq \tilde{C}_\pi r^{\frac{1}{2}\lfloor \frac{t}{2d'} \rfloor} (1 + 3(a_2 + a_3 + a_4 + a_5)) + \left(\frac{\tilde{C}_J}{\eta} + 1\right) \beta^{\lfloor \frac{t}{2} \rfloor + 3}
 \end{aligned}$$



where  $\tilde{C}_\pi := \int \left( \frac{\max_i \{1, v_i\}}{\min_i \{1, v_i\}} \right) \left( \prod_{i=1}^4 v_i^{a_i + a_{i+1} - 1} \right) \exp \left( \sum_{i=1}^5 -v_i v_{i-1} \right) dv / C_g$  and

$\tilde{C}_J := \frac{1}{C_g} \int J_0 \left( \prod_{i=1}^4 v_i^{a_i + a_{i+1} - 1} \right) \exp \left( \sum_{i=1}^5 -v_i v_{i-1} \right) dv$ , and we derive a bound for these terms in the Appendix.

For the purpose of illustrating this result in a concrete example, let us set  $x = 2$ ,  $b = 3$  and  $a_i = i$ . Then by (2.60)  $\tilde{C}_\pi \leq 60,300$  and by (2.61)  $\tilde{C}_J \leq 59$ ,  $\beta \leq 7/9$ ,  $r \leq 1 - \frac{1}{2178}$ ,  $20 \leq \eta \leq 21$  and  $216 \leq d' \leq 217$ , and hence

$$d_{TV}(\mathcal{U}^{t+3}, \pi) \leq 60300 * 43 \left( 1 - \frac{1}{2178} \right)^{\frac{1}{2} \lfloor \frac{t}{434} \rfloor} + \left( 1 + \frac{59}{20} \right) \left( \frac{7}{9} \right)^{\lfloor \frac{t}{2} \rfloor + 3}$$

which implies that  $d_{TV}(\mathcal{U}^{t+3}, \pi) \leq 10^{-5}$  for  $t \geq 50,000,000$ .

**2.2. A brief look at the case  $n = 3$ .** The case  $n = 3$  can be treated in a very similar manner as was done for  $n = 4$ . It follows immediately from (1.4) that this problem would reduce to dealing with a Markov chain of a single variable, given by

$$u^{t+1} = \frac{\gamma_2^{t+1}}{\frac{\gamma_1^{t+1}}{u^t + x} + \frac{\gamma_3^{t+1}}{u^t + b}}$$

Similarly, coupling two copies  $(u^t, v^t)$  uniformly with the property  $u^0 \leq v^0$  implies that  $u^t \leq v^t$ . It is also an immediate observation that the ratio  $R_t = \frac{v^t}{u^t}$  is strictly decreasing, hence we no longer need to define a process like (2.1) and can simply work with this ratio directly. It follows that  $R_{t+1} = R_t Q_t$  where

$$\begin{aligned} Q_t &:= 1 - \frac{\left( 1 - \frac{1}{R_t} \right) \left( x \gamma_1^{t+1} / ((u^t + x) (1 + \frac{x}{v^t})) + b \gamma_3^{t+1} / ((u^t + b) (1 + \frac{b}{v^t})) \right)}{\left( \gamma_1^{t+1} / (1 + \frac{x}{v^t}) + \gamma_3^{t+1} / (1 + \frac{b}{v^t}) \right)} \\ &\leq 1 - \frac{\left( 1 - \frac{1}{R_t} \right) (x \gamma_1^{t+1} + b \gamma_3^{t+1})}{\left( \gamma_1^{t+1} / (1 + \frac{x}{v^t}) + \gamma_3^{t+1} / (1 + \frac{b}{v^t}) \right) (u^t + \max\{x, b\}) \left( 1 + \frac{\max\{x, b\}}{u^t} \right)} \\ &\leq \dot{r}_t + \frac{1 - \dot{r}_t}{R_t} \end{aligned}$$

where  $\dot{r}_t := 1 - \min\{x, b\} / ((u^t + \max\{x, b\}) (1 + \frac{\max\{x, b\}}{u^t}))$ . Note that if we define  $K_{1,t+1} := \frac{\gamma_2^{t+1}}{\gamma_1^{t+1} + \gamma_3^{t+1}} (u^t + x + b)$  and  $K_{2,t+1} := \left( \frac{\gamma_1^{t+1}}{\gamma_2^{t+1}} \frac{1}{x} + \frac{\gamma_3^{t+1}}{\gamma_2^{t+1}} \frac{1}{b} \right)$  then  $K_{1,t+1} \geq u^{t+1}$  and  $K_{2,t+1} \geq \frac{1}{u^{t+1}}$ , and hence we do not need a process analogous to  $D_t$  from the previous section, since

$$\dot{r}_{t+1} \leq 1 - \min\{x, b\} / ((K_{1,t+1} + \max\{x, b\}) (1 + \max\{x, b\} K_{2,t+1}))$$

If the process  $J_t$  and the stopping time  $S$ , as well as the constant  $\eta$  are also defined in an analogous manner, then we can repeat the steps of (2.9)

$$\begin{aligned} \mathbb{E}[R_{S+1}] &= \mathbb{E}[Q_S R_S] \\ &\leq \mathbb{E}[\dot{r}_S (R_S - 1)] + 1 \\ &= \mathbb{E} \left[ \left( 1 - \frac{\min\{x, b\}}{(u^S + \max\{x, b\}) \left( 1 + \frac{\max\{x, b\}}{u^S} \right)} \right) (R_S - 1) \right] + 1 \\ &\leq \mathbb{E} \left[ \left( 1 - \frac{\min\{x, b\}}{(\eta + \max\{x, b\}) (1 + \eta \max\{x, b\})} \right) (R_S - 1) \right] + 1 \\ (2.27) \quad &= r \mathbb{E}[R_S - 1] + 1 \end{aligned}$$

where  $r = 1 - \min\{x, b\} / ((\eta + \max\{x, b\})(1 + \eta \max\{x, b\}))$ . Note that we no longer need to look at time  $S + 2$  in the left-hand side of (2.27) in order to obtain this inequality. This means that from the proof of Lemma 12 and Corollary 13 we get

$$\mathbb{E}[R_{t+1} | J_0 \leq \eta] \leq 1 + 3r^{\lfloor \frac{t}{d'} \rfloor} (R_0 - 1)$$

From the proof of Theorem 1 we conclude

$$d_{TV}(u^{t+2}, v^{t+2}) \leq r^{\lfloor \frac{t}{2d'} \rfloor} (1 + 3(a_2 + a_3)(R_0 - 1)) + \frac{\max\{J_0, \eta\} \beta^{\lfloor \frac{t}{2} \rfloor + 3}}{\eta}$$

We can make an analogous argument to obtain a result similar to Corollary 2. In particular if we let  $\mathcal{U}^0 = (1, 1, 1)$ ,  $\mathcal{V}^0 \sim \pi$  and  $x = 1$ ,  $b = 2$  and  $a_i = i$ , then by calculations similar to those done in Section 2.1.3 we get

$$d_{TV}(\mathcal{U}^{t+2}, \pi) \leq 600 \left(1 - \frac{1}{65}\right)^{\lfloor \frac{t}{100} \rfloor} + 6 \left(\frac{7}{9}\right)^{\lfloor \frac{t}{2} \rfloor + 3}$$

which in particular implies that  $d_{TV}(\mathcal{U}^{t+2}, \pi) \leq 10^{-5}$  for  $t \geq 125,000$ .

**2.3.  $n > 4$ .** It is not difficult to show that the method outlined in (2.1) and leading to inequality (2.2) can also be extended to the case where  $n = 5$ : given two starting points  $u^0 \preceq v^0 \in \mathbb{R}_{+}^5$ , it amounts to setting  $u^{t+1} := f^{t+1}(u^t)$  and  $v^{t+1} := \lambda_{t+1} u^{t+1}$ , where  $\lambda_{t+1} := \max\left\{\frac{f_2^{t+1}(v^t)}{u_2^{t+1}}, \frac{f_4^{t+1}(v^t)}{u_4^{t+1}}\right\}$ . A calculation similar to (2.2) shows that  $R_{t+1} = \lambda_{t+1} < R_t$  whenever  $u^t \neq v^t$ . If we attempt to replicate this method for  $n \geq 6$  however, it becomes apparent that  $\lambda_t = R_0$ , so that  $R_t = R_0$  is fixed for all times  $t$ . It is nonetheless possible to extend this method for  $n \geq 6$  if we consider a multi-step version of the Markov chain in (1.4). More precisely, let  $f_m : \mathbb{R}_{+}^m \rightarrow \mathbb{R}_{+}^m$  be defined by

$$\begin{aligned} f_m(u_2, u_4, \dots, u_{2m}) &:= \left( \frac{\gamma_2}{\frac{\gamma_1}{x+u_2} + \frac{\gamma_3}{u_2+u_4}}, \frac{\gamma_4}{\frac{\gamma_3}{u_2+u_4} + \frac{\gamma_5}{u_4+u_6}}, \dots, \frac{\gamma_{2m}}{\frac{\gamma_{2m-1}}{u_{2(m-1)}+u_{2m}} + \frac{\gamma_{2m+1}}{u_{2m}+b}} \right) \\ &= (f_{(m,2)}(u), f_{(m,4)}(u), \dots, f_{(m,2m)}(u)) \end{aligned}$$

and let

$$(2.28) \quad F_m^k := f_m^k \circ f_m^{k-1} \circ \dots \circ f_m^1$$

We can show that in this case  $R_t$  is strictly decreasing. However, it is not a straight forward matter to show (and derive an upper bound for) the geometric convergence of  $R_t \rightarrow 1$ , primarily due to the difficulty of handling the 'continued fraction' form of the functions  $h_t$ . We will therefore consider an alternative Markov chain. For odd  $n = 2m + 1$  we will define this by

$$(2.29) \quad (u_2^{t+1}, u_4^{t+1}, \dots, u_{2m}^{t+1}) = g(u_2^t, u_4^t, \dots, u_{2m}^t) := \left( \frac{\gamma_2}{\frac{\tilde{\gamma}_1}{x+u_2^t} + \frac{\gamma_3}{u_2^t+u_4^t}}, \frac{\gamma_4}{\frac{\tilde{\gamma}_3}{u_2^{t+1}+u_4^t} + \frac{\gamma_5}{u_4^t+u_6^t}}, \dots, \frac{\gamma_{2m}}{\frac{\tilde{\gamma}_{2m-1}}{u_{2m-2}^{t+1}+u_{2m}^t} + \frac{\gamma_{2m+1}}{u_{2m}^t+b}} \right)$$

where  $\{\gamma_2, \gamma_3, \dots, \gamma_{2n+1}\}$  are same as before and  $\{\tilde{\gamma}_1, \tilde{\gamma}_3, \dots, \tilde{\gamma}_{2n-1}\}$  is an i.i.d. copy of  $\{\gamma_1, \gamma_3, \dots, \gamma_{2n-1}\}$ . The definition of  $g$  for  $n$  even is as one might expect. Furthermore it will be apparent that all arguments for  $n$  even would be non-distinct from ones about to be made for odd  $n$ , which is why we shall forgo the separate treatment of this case.

Note also that the random function in (2.29) can be identified with the transition kernel

$$(2.30) \quad \bar{O} := (\bar{P}_1 \bar{P}_3 \bar{P}_5 \dots \bar{P}_{2m+1}) (\bar{P}_2 \bar{P}_3 \bar{P}_4 \dots \bar{P}_{2m})$$

We observe that the kernel  $(\bar{P}_1 \bar{P}_3 \bar{P}_5 \dots \bar{P}_{2m+1})$  is responsible for generating the variables  $\{\tilde{\gamma}_1, \gamma_3, \gamma_5, \dots, \gamma_{2m+1}\}$  while  $(\bar{P}_2 \bar{P}_3 \bar{P}_4 \dots \bar{P}_{2m})$  generates  $\{\gamma_2, \tilde{\gamma}_3, \gamma_4, \tilde{\gamma}_5, \dots, \tilde{\gamma}_{2m-1}, \gamma_{2m}\}$ . Recalling that  $\pi \bar{P}_i = \pi$ , it is evident that  $\pi$  is also invariant with respect to  $\bar{O}$ , which shows that this Markov chain will also converge to  $\pi$  in distribution (it is not difficult to ascertain that this is indeed a Harris chain).

We can now extend the method we used for the case  $n = 4$  to general  $n$ , with the Markov chain defined by the random functions  $\{g^t\}$ : starting with  $u^0 \preceq v^0$  such that  $\frac{v_2^0}{u_2^0} = \dots = \frac{v_{2m}^0}{u_{2m}^0} > 1$ , set  $u^{t+1} := g^{t+1}(u^t)$  and  $v^{t+1} := R_{t+1}u^{t+1}$  where

$$(2.31) \quad R_{t+1} := \max_j \left\{ \frac{g_{2j}^{t+1}(v^t)}{g_{2j}^{t+1}(u^t)} \right\} = \frac{v_2^t}{u_2^t} \max \left\{ \frac{\frac{\tilde{\gamma}_1^{t+1}}{1+\frac{x}{u_2^t}} + \frac{\gamma_3^{t+1}}{1+\frac{u_4^t}{u_2^t}}}{\frac{\tilde{\gamma}_1^{t+1}}{1+\frac{x}{v_2^t}} + \frac{\gamma_3^{t+1}}{1+\frac{v_4^t}{v_2^t}}}, \frac{\frac{\tilde{\gamma}_3^{t+1}}{1+\frac{g_{2j}^{t+1}(u^t)}{u_4^t}} + \frac{\gamma_5^{t+1}}{1+\frac{u_6^t}{u_4^t}}}{\frac{\tilde{\gamma}_3^{t+1}}{1+\frac{g_{2j}^{t+1}(v^t)}{v_4^t}} + \frac{\gamma_5^{t+1}}{1+\frac{v_6^t}{v_4^t}}}, \dots, \frac{\frac{\tilde{\gamma}_{2m-1}^{t+1}}{1+\frac{g_{2m-2}^{t+1}(u^t)}{u_{2m}^t}} + \frac{\gamma_{2m+1}^{t+1}}{1+\frac{b}{u_{2m}^t}}}{\frac{\tilde{\gamma}_{2m-1}^{t+1}}{1+\frac{g_{2m-2}^{t+1}(v^t)}{v_{2m}^t}} + \frac{\gamma_{2m+1}^{t+1}}{1+\frac{b}{v_{2m}^t}}} \right\}$$

Here we have used the notation  $g = (g_2, g_4, \dots, g_{2m})$  to represent the components of the function  $g$ . In (2.31) we have extracted the factor  $\frac{v_2^t}{u_2^t}$  by making the implicit assumption that  $\frac{v_2^t}{u_2^t} = \frac{v_4^t}{u_4^t} = \dots = \frac{v_{2m}^t}{u_{2m}^t}$ . The validity of this is evident from the definition of the process  $v^t$ , and inductively (in  $t$ ) from (2.31). Let  $R_0 = \frac{v_2^0}{u_2^0}$ . We can then confirm by a simple inductive argument that  $\frac{g_{2j}^{t+1}(v^t)}{g_{2j}^{t+1}(u^t)} < R_t$  for all  $j$  and  $t \geq 0$  as follows: it is

$$(2.32) \quad \text{immediate that } \frac{g_{2j}^{t+1}(v^t)}{g_{2j}^{t+1}(u^t)} = R_t \left( \frac{\frac{\tilde{\gamma}_1^{t+1}}{1+\frac{x}{u_2^t}} + \frac{\gamma_3^{t+1}}{1+\frac{u_4^t}{u_2^t}}}{\frac{\tilde{\gamma}_1^{t+1}}{1+\frac{x}{v_2^t}} + \frac{\gamma_3^{t+1}}{1+\frac{v_4^t}{v_2^t}}} \right) < R_t, \text{ since } \frac{u_4^t}{u_2^t} = \frac{v_4^t}{v_2^t} \text{ while } \frac{x}{u_2^t} > \frac{x}{v_2^t}. \text{ Now assuming that } \frac{g_{2j}^{t+1}(v^t)}{g_{2j}^{t+1}(u^t)} < R_t, \text{ we get}$$

$$\frac{g_{2j}^{t+1}(u^t)}{u_{2j+2}^t} > \frac{g_{2j}^{t+1}(v^t)}{v_{2j+2}^t}$$

$$(2.33) \quad \text{(since } \frac{v_{2j+2}^t}{u_{2j+2}^t} = R_t \text{), hence } \frac{g_{2j+2}^{t+1}(v^t)}{g_{2j+2}^{t+1}(u^t)} = R_t \left( \frac{\frac{\tilde{\gamma}_{2j+1}^{t+1}}{1+\frac{g_{2j}^{t+1}(u^t)}{u_{2j+2}^t}} + \frac{\gamma_{2j+3}^{t+1}}{1+\frac{u_{2j+4}^t}{u_{2j+2}^t}}}{\frac{\tilde{\gamma}_{2j+1}^{t+1}}{1+\frac{g_{2j}^{t+1}(v^t)}{v_{2j+2}^t}} + \frac{\gamma_{2j+3}^{t+1}}{1+\frac{v_{2j+4}^t}{v_{2j+2}^t}}} \right) < R_t \text{ (again since } \frac{u_{2j+4}^t}{u_{2j+2}^t} \geq \frac{v_{2j+4}^t}{v_{2j+2}^t} \text{, where by convention we take } u_{2m+2}^t = v_{2m+2}^t = b \text{), which completes this inductive argument.}$$

Let us now consider the  $i^{th}$  term in the right-hand side of (2.31). If we replace both  $\gamma_{2i+1}^{t+1} / (1 + u_{2i+2}^t / u_{2i}^t)$  in the numerator and  $\gamma_{2i+1}^{t+1} / (1 + v_{2i+2}^t / v_{2i}^t)$  in the denominator by  $\gamma_{2i+1}^{t+1}$ , then by Lemma 7 the right-hand side of (2.31) would not decrease. Hence we can say that

$$(2.33) \quad R_{t+1} \leq \frac{v_2^t}{u_2^t} \max \left\{ \frac{\frac{\tilde{\gamma}_1^{t+1}}{1+\frac{x}{u_2^t}} + \gamma_3^{t+1}}{\frac{\tilde{\gamma}_1^{t+1}}{1+\frac{x}{v_2^t}} + \gamma_3^{t+1}}, \frac{\frac{\tilde{\gamma}_3^{t+1}}{1+\frac{g_{2j}^{t+1}(u^t)}{u_4^t}} + \gamma_5^{t+1}}{\frac{\tilde{\gamma}_3^{t+1}}{1+\frac{g_{2j}^{t+1}(v^t)}{v_4^t}} + \gamma_5^{t+1}}, \dots, \frac{\frac{\tilde{\gamma}_{2m-1}^{t+1}}{1+\frac{g_{2m-2}^{t+1}(u^t)}{u_{2m}^t}} + \gamma_{2m+1}^{t+1}}{\frac{\tilde{\gamma}_{2m-1}^{t+1}}{1+\frac{g_{2m-2}^{t+1}(v^t)}{v_{2m}^t}} + \gamma_{2m+1}^{t+1}} \right\} =: \frac{v_2^t}{u_2^t} Q_{t+1}$$

We can proceed in a manner similar to what we did in Lemma 8. Let  $\mu_1 := a_3 + a_4 + \dots + a_{2m+2}$  and  $\mu_2 := \frac{1}{a_1+a_2-1} + \dots + \frac{1}{a_{2m-1}+a_{2m}-1}$ . Then by [6] and Markov's inequality

$$(2.34) \quad \begin{aligned} \mathbb{P} \left[ \cap_{i \text{ odd}, i=1}^{2m-1} \left\{ \tilde{\gamma}_i \geq \frac{1}{2\mu_2} \right\}, \cap_{i \text{ odd}, i=3}^{2m+1} \{ \gamma_i \leq \mu_1 \} \right] &= \mathbb{P} \left[ \cap_{i \text{ odd}, i=1}^{2m-1} \left\{ \frac{1}{\tilde{\gamma}_i} < 2\mu_2 \right\} \right] \mathbb{P} \left[ \cap_{i \text{ odd}, i=3}^{2m+1} \{ \gamma_i \leq \mu_1 \} \right] \\ &\geq \mathbb{P} \left[ \sum_{i \text{ odd}, i=1}^{2m-1} \frac{1}{\tilde{\gamma}_i} < 2\mu_2 \right] \mathbb{P} \left[ \sum_{i \text{ odd}, i=3}^{2m+1} \gamma_i \leq \mu_1 \right] \\ &\geq \frac{1}{4} \end{aligned}$$

**Lemma 14.** Let  $M_{t+1} := \max_{2 \leq i \leq m} \left\{ 2 + \frac{g_{2i-2}^{t+1}(v^t)}{v_{2i}^t} + \frac{u_{2i}^t}{u_{2i-2}^{t+1}} \right\}$  and  $\dot{r}_{t+1} := 1 - \frac{1}{(1+2\mu_1\mu_2)^m (M_{t+1})^{m-1} \left(1 + \frac{x}{v_2^t}\right) \left(1 + \frac{u_2^t}{x}\right)}$ . Then

$$\mathbb{E}[R_{t+1}] \leq \mathbb{E} \left[ \left( \frac{1}{4} \dot{r}_{t+1} + \frac{3}{4} \right) (R_t - 1) \right] + 1$$

*Proof.* Let  $\dot{r}_{t+1,2} := 1 - \frac{1}{(1+2\mu_1\mu_2) \left(1 + \frac{x}{v_2^t}\right) \left(1 + \frac{u_2^t}{x}\right)}$ , and define recursively  $\dot{r}_{t+1,2i} := 1 - \frac{(1-\dot{r}_{t+1,2i-2})}{(1+2\mu_1\mu_2)M_{t+1}}$  for  $2 \leq i \leq m$ . Let  $Q_{t+1,i}$  be the  $i^{th}$  term inside the max in (2.33). We claim that in the event  $\left\{ \cap_{i \text{ odd}, i=1}^{2m-1} \left\{ \gamma_i^{t+1} \geq \frac{1}{2\mu_2} \right\}, \cap_{i \text{ odd}, i=3}^{2m+1} \left\{ \gamma_i^{t+1} \leq \mu_1 \right\} \right\}$ , the term  $Q_{t+1,i}$  is bounded from above by  $\dot{r}_{t+1,2i} + \frac{1-\dot{r}_{t+1,2i}}{R_t}$  for each  $i$ . Assume that this statement is true for  $i-1$ . Note that since the  $i-1^{st}$  term inside the  $\max$  in (2.31) is less than or equal to the  $i-1^{st}$  term inside the  $\max$  in (2.33), this implies that  $\frac{g_{2i-2}^{t+1}(v^t)}{g_{2i-2}^{t+1}(u^t)} \leq \frac{v_2^t}{u_2^t} Q_{t+1,i-1} \leq R_t \left( \dot{r}_{t+1,2i-2} + \frac{1-\dot{r}_{t+1,2i-2}}{R_t} \right)$ . Then by Lemma 7 and the fact that  $\frac{g_{2i-2}^{t+1}(u^t)}{u_{2i}^t} > \frac{g_{2i-2}^{t+1}(v^t)}{v_{2i}^t}$  (which follows from (2.32)), we get

$$\begin{aligned} \frac{\frac{\tilde{\gamma}_{2i-1}^{t+1}}{1 + \frac{u_{2i}^t}{g_{2i-2}^{t+1}(u^t)}} + \gamma_{2i+1}^{t+1}}{\frac{\tilde{\gamma}_{2i-1}^{t+1}}{1 + \frac{g_{2i-2}^{t+1}(v^t)}{v_{2i}^t}} + \gamma_{2i+1}^{t+1}} &\leq \frac{\frac{\frac{1}{2\mu_2}}{1 + \frac{g_{2i-2}^{t+1}(u^t)}{v_{2i}^t}} + \mu_1}{\frac{\frac{1}{2\mu_2}}{1 + \frac{g_{2i-2}^{t+1}(v^t)}{v_{2i}^t}} + \mu_1} = 1 - \frac{\frac{1}{1 + \frac{g_{2i-2}^{t+1}(v^t)}{v_{2i}^t}} - \frac{1}{1 + \frac{g_{2i-2}^{t+1}(u^t)}{v_{2i}^t}}}{\frac{1}{1 + \frac{g_{2i-2}^{t+1}(v^t)}{v_{2i}^t}} + 2\mu_1\mu_2} \\ &= 1 - \frac{1 - \frac{g_{2i-2}^{t+1}(v^t)}{v_{2i}^t} \frac{u_{2i}^t}{g_{2i-2}^{t+1}(u^t)}}{\left( \frac{1}{1 + \frac{g_{2i-2}^{t+1}(v^t)}{v_{2i}^t}} + 2\mu_1\mu_2 \right) \left( 1 + \frac{g_{2i-2}^{t+1}(v^t)}{v_{2i}^t} \right) \left( 1 + \frac{g_{2i-2}^{t+1}(u^t)}{u_{2i}^t} \right) \frac{u_{2i}^t}{g_{2i-2}^{t+1}(u^t)}} \\ &\leq 1 - \frac{1 - \frac{1}{R_t} \left( R_t \left( \dot{r}_{t+1,2i-2} + \frac{1-\dot{r}_{t+1,2i-2}}{R_t} \right) \right)}{\left( 1 + 2\mu_1\mu_2 \right) \left( 1 + \frac{g_{2i-2}^{t+1}(v^t)}{v_{2i}^t} \right) \left( 1 + \frac{u_{2i}^t}{g_{2i-2}^{t+1}(u^t)} \right)} \\ &\leq 1 - \frac{(1 - \dot{r}_{t+1,2i-2}) \left( 1 - \frac{1}{R_t} \right)}{\left( 1 + 2\mu_1\mu_2 \right) \left( 2 + \frac{g_{2i-2}^{t+1}(v^t)}{v_{2i}^t} + \frac{u_{2i}^t}{g_{2i-2}^{t+1}(u^t)} \right)} \\ &\leq \dot{r}_{t+1,2i} + \frac{1 - \dot{r}_{t+1,2i}}{R_t} \end{aligned}$$

This proves the inductive step. The computation showing  $Q_{t+1,1} \leq \dot{r}_{t+1,2} + \frac{1-\dot{r}_{t+1,2}}{R_t}$  is identical to the one given above. Observe that  $\dot{r}_{t+1,2} \leq \dot{r}_{t+1,4} \leq \dots \leq \dot{r}_{t+1,2m}$ , or more precisely

$$\begin{aligned} 1 - \dot{r}_{t+1,2i} &= \frac{(1 - \dot{r}_{t+1,2i-2})}{(1 + 2\mu_1\mu_2) M_{t+1}} \\ (2.35) \quad &= \left( \frac{1}{(1 + 2\mu_1\mu_2) M_{t+1}} \right)^{i-1} (1 - \dot{r}_{t+1,2}) \\ &= \frac{1}{(1 + 2\mu_1\mu_2)^i (M_{t+1})^{i-1} \left( 1 + \frac{x}{v_2^t} \right) \left( 1 + \frac{u_2^t}{x} \right)} \end{aligned}$$

We conclude that  $Q_{t+1} \leq \dot{r}_{t+1,2m} + \frac{1-\dot{r}_{t+1,2m}}{R_t}$ , and hence by (2.34)

$$\mathbb{E}[R_{t+1}] \leq \mathbb{E}\left[\left(\frac{1}{4}\dot{r}_{t+1} + \frac{3}{4}\right)(R_t - 1)\right] + 1$$

□

An apparent weakness of (2.35) is that the bound on the rate of convergence is exponentially bad in  $n$ . Unfortunately, this is an inherent property of this method: observe that even under the most favourable scenario in (2.31) (whereby we take  $\gamma_3 = \gamma_5 = \dots = \gamma_{2n+1} = 0$ ), one still arrives at

$$\begin{aligned} \dot{r}_{t+1,2} &= 1 - \frac{1}{1 + \frac{u_2^t}{x}} \\ \dot{r}_{t+1,2i} &= 1 - \frac{(1 - \dot{r}_{t+1,2i-2})}{1 + \frac{u_{2i-2}^t}{u_{2i-2}^{t+1}}} \\ &\geq 1 - \left(\frac{1}{m_{t+1}}\right)^i \end{aligned}$$

where  $m_{t+1} = \min_i \left\{1 + \frac{u_{2i}^t}{u_{2i-2}^{t+1}}\right\}$ .

Note that the results derived in section 2.1.1 are in fact independent of many aspects of the model, including the dimension  $n$  as well as the Markov chain in question. They relied only on the existence of auxiliary processes  $(K_{i,t}, J_t, D_t)$ , stochastic monotonicity of the two paths, and the existence of a non-increasing process  $R_t$  as described in the prelude to Lemma 6. At this point we only require to ascertain the existence of an auxiliary processes  $\{K_t\}$  that satisfies conditions similar to the ones we had for the case  $n = 4$ , and which assists in bounding  $M_t$  from above. We will prove such existence under certain restrictions.

**2.3.1. Construction of  $\{K_{i,t}\}$  for  $n \geq 5$ .** For  $i = 1, \dots, m$  let  $w_{2i}^t := \frac{1}{u_{2i}^t}$ . Then

$$w_{2i}^{t+1} = \frac{1}{\gamma_{2i}^{t+1}} \left( \frac{\tilde{\gamma}_{2i-1}^{t+1}}{\frac{1}{w_{2i-2}^{t+1}} + \frac{1}{w_{2i}^t}} + \frac{\gamma_{2i+1}^{t+1}}{\frac{1}{w_{2i}^t} + \frac{1}{w_{2i+2}^t}} \right)$$

where for convenience we have taken  $w_0^t = \frac{1}{x}$  and  $w_{2m+2}^t = \frac{1}{b}$  for all  $t$ . We will make use of the following inequality: for  $a, b, \rho_1, \rho_2 \in \mathbb{R}^+$ ,

$$(2.36) \quad \left(\frac{a}{\rho_1} + \frac{b}{\rho_2}\right) \geq \frac{(a+b)^2}{(a\rho_1 + b\rho_2)}$$

Hence, for  $1 \leq i \leq m$

$$\begin{aligned} w_{2i}^{t+1} &= \frac{1}{\gamma_{2i}^{t+1}} \left( \frac{\tilde{\gamma}_{2i-1}^{t+1}}{\frac{1}{w_{2i-2}^{t+1}} + \frac{1}{w_{2i}^t}} + \frac{\gamma_{2i+1}^{t+1}}{\frac{1}{w_{2i}^t} + \frac{1}{w_{2i+2}^t}} \right) \\ (2.37) \quad &\leq \frac{1}{\gamma_{2i}^{t+1}} \left( \frac{\tilde{\gamma}_{2i-1}^{t+1}}{4} (w_{2i-2}^{t+1} + w_{2i}^t) + \frac{\gamma_{2i+1}^{t+1}}{4} (w_{2i}^t + w_{2i+2}^t) \right). \end{aligned}$$

We can now exploit the linearity in (2.37) to get an upper bound on  $\mathbb{E}[K_{1,t+1} | \mathcal{F}_t]$  where  $K_{1,t} := \sum_{i=1}^m w_{2i}^t$ . Observe that

$$\begin{aligned} \mathbb{E}[K_{1,t+1} | \mathcal{F}_t] &= \sum_{i=1}^m \mathbb{E}[w_{2i}^{t+1} | \mathcal{F}_t] \\ (2.38) \quad &\leq \sum_{i=1}^m \frac{1}{\alpha_{2i} - 1} \left( \frac{\alpha_{2i-1}}{4} (\mathbb{E}[w_{2i-2}^{t+1} | \mathcal{F}_t] + w_{2i}^t) + \frac{\alpha_{2i+1}}{4} (w_{2i}^t + w_{2i+2}^t) \right) \end{aligned}$$

We will re-write the right-hand side of (2.38) in a form that will reduce it to a super-martingale type of inequality, analogous to (2.22) for the  $n = 4$  case. Let  $A_i = \mathbb{E}[w_{2i}^{t+1} | \mathcal{F}_t]$ ,  $B_i = w_{2i}^t$  for  $1 \leq i \leq m$ , and  $A_0 = B_0 = \frac{1}{x}$  and  $A_{m+1} = B_{m+1} = \frac{1}{b}$ . Let  $C_i^+ = \frac{\alpha_{2i+1}}{4(\alpha_{2i}-1)}$ ,  $C_i^- = \frac{\alpha_{2i-1}}{4(\alpha_{2i}-1)}$  and  $D_i = C_i^+ + C_i^-$  for  $1 \leq i \leq m$ , and  $C_0^+ = C_0^- = 0$ . Then by (2.37), for  $1 \leq i \leq m$

$$(2.39) \quad A_i \leq C_i^- A_{i-1} + D_i B_i + C_i^+ B_{i+1}$$

In particular, since  $A_0 = \frac{1}{x} = B_0$

$$(2.40) \quad A_1 \leq C_1^- B_0 + D_1 B_1 + C_1^+ B_2$$

Define  $q_{i,j}$  as follows:  $q_{1,0} = C_1^-$ ,  $q_{1,1} = D_1$ ,  $q_{1,2} = C_1^+$ , and  $q_{1,j} = 0$  for  $j > 2$  (so that  $A_1 \leq \sum_{j=0}^{m+1} q_{1,j} B_j$  by (2.40)); and for  $2 \leq i \leq m$ ,

$$\begin{aligned} q_{i,i+1} &= C_i^+ \\ q_{i,i} &= C_i^- q_{i-1,i} + D_i \\ q_{i,j} &= C_i^- q_{i-1,j} \quad \text{for } 0 \leq j < i \\ q_{i,j} &= 0 \quad \text{for } j > i+1 \end{aligned}$$

Then for  $2 \leq i \leq m$

$$A_i \leq \sum_{j=0}^{m+1} q_{i,j} B_j$$

which follows from (2.39) and (2.40) and by induction on  $i$ .

Then the formulas for  $q_{i,j}$  ( $1 \leq i \leq m$ ) are:

$$q_{i,0} = \prod_{k=1}^i C_k^-$$

and for  $1 \leq j \leq m+1$

$$\begin{aligned} q_{i,j} &= 0 \quad \text{for } i < j-1 \\ q_{j-1,j} &= C_{j-1}^+ \\ q_{j,j} &= C_j^- q_{j-1,j} + D_j = C_j^- C_{j-1}^+ + D_j \\ q_{i,j} &= C_i^- q_{i-1,j} = (C_j^- C_{j-1}^+ + D_j) \prod_{k=j+1}^i C_k^- \quad \text{for } i > j \end{aligned}$$

Next, let

$$(2.41) \quad \zeta_{2j} = \sum_{i=1}^m q_{i,j}$$

Then

$$(2.42) \quad \mathbb{E}[K_{1,t+1} | \mathcal{F}_t] = \sum_{i=1}^m A_i \leq \sum_{j=0}^{m+1} \zeta_{2j} B_j = \sum_{j=0}^{m+1} \zeta_{2j} w_{2i}^t$$

We have  $\zeta_{2m+2} = q_{m,m+1} = C_m^+$  and  $\zeta_0 = \sum_{i=1}^m \prod_{k=1}^i C_k^-$ . For  $1 \leq j \leq m$ , we have

$$\begin{aligned} \zeta_{2j} &= \sum_{i=j-1}^m q_{i,j} \\ (2.43) \quad &= C_{j-1}^+ + (C_j^- C_{j-1}^+ + D_j) \left( \sum_{i=j}^m \prod_{k=j+1}^i C_k^- \right) \end{aligned}$$

We can obtain a similar result for  $\mathbb{E}[K_{2,t} | \mathcal{F}_t]$  where  $K_{2,t} := \sum_{i=1}^m u_{2i}^t$ . Setting  $u_0^t = x$  and  $u_{n+1}^t = b$  for all  $t$ , it follows from (2.36) that for  $1 \leq i \leq m$

$$\begin{aligned} u_{2i}^{t+1} &= \frac{\gamma_{2i}^{t+1}}{\frac{\tilde{\gamma}_{2i-1}^{t+1}}{u_{2i-2}^{t+1} + u_{2i}^t} + \frac{\gamma_{2i+1}^{t+1}}{u_{2i}^t + u_{2i+2}^t}} \\ &\leq \gamma_{2i}^{t+1} \left( \frac{\tilde{\gamma}_{2i-1}^{t+1}}{(\tilde{\gamma}_{2i-1}^{t+1} + \gamma_{2i+1}^{t+1})^2} (u_{2i-2}^{t+1} + u_{2i}^t) + \frac{\gamma_{2i+1}^{t+1}}{(\tilde{\gamma}_{2i-1}^{t+1} + \gamma_{2i+1}^{t+1})^2} (u_{2i}^t + u_{2i+2}^t) \right) \end{aligned}$$

Since  $\frac{1}{(\tilde{\gamma}_{2i-1}^{t+1} + \gamma_{2i+1}^{t+1})^2}$  is a decreasing function of  $\tilde{\gamma}_{2i-1}^{t+1}$ , by Harris' inequality (page 136 of [2]) we get

$$\begin{aligned} \mathbb{E} \left[ \frac{\tilde{\gamma}_{2i-1}^{t+1}}{(\tilde{\gamma}_{2i-1}^{t+1} + \gamma_{2i+1}^{t+1})^2} \right] &\leq \frac{\alpha_{2i-1}}{(\alpha_{2i-1} + \alpha_{2i+1} - 1)(\alpha_{2i-1} + \alpha_{2i+1} - 2)}. \end{aligned}$$

Therefore, in an analogous manner to the previous derivations, we let  $E_i = \mathbb{E}[u_{2i}^{t+1} | \mathcal{F}_t]$ ,  $F_i = u_{2i}^t$  for  $1 \leq i \leq m$ , and  $E_0 = F_0 = x$  and  $E_{m+1} = F_{m+1} = b$ . Let  $G_i^+ = \frac{\alpha_{2i} \alpha_{2i+1}}{(\alpha_{2i-1} + \alpha_{2i+1} - 1)(\alpha_{2i-1} + \alpha_{2i+1} - 2)}$ ,  $G_i^- = \frac{\alpha_{2i} \alpha_{2i-1}}{(\alpha_{2i-1} + \alpha_{2i+1} - 1)(\alpha_{2i-1} + \alpha_{2i+1} - 2)}$  and  $H_i = G_i^+ + G_i^-$  for  $1 \leq i \leq m$ , and  $G_0^+ = G_0^- = 0$ . We also define  $p_{i,j}$  in an analogous manner to  $q_{i,j}$  so that  $E_i \leq \sum_{j=0}^{m+1} p_{ij} F_j$ , and let

$$\begin{aligned} \xi_{2j} &= \sum_{i=1}^m p_{i,j} \\ (2.44) \quad &= G_{j-1}^+ + (G_j^- G_{j-1}^+ + D_j) \left( \sum_{i=j}^m \prod_{k=j+1}^i H_k^- \right) \end{aligned}$$

Then

$$(2.45) \quad \mathbb{E}[K_{2,t+1} | \mathcal{F}_t] = \sum_{i=1}^m E_i \leq \sum_{j=0}^{m+1} \xi_{2j} F_j = \sum_{j=0}^{m+1} \xi_{2j} u_{2j}^t$$

It is now immediate that  $\mathbb{E}[K_{2,t+1} | \mathcal{F}_t] \leq \max\{\xi_i\} K_{2,t} + C_2$  and  $\mathbb{E}[K_{1,t+1} | \mathcal{F}_t] \leq \max\{\zeta_i\} K_{1,t} + C_1$  where  $C_1 := \zeta_0 \frac{1}{x} + \zeta_{2m+2} \frac{1}{m}$  and  $C_2 := \xi_0 x + \xi_{2m+2} m$ , which was the goal of the last derivations.

We can now repeat the argument that led to (2.9), to obtain the following analogous inequality whenever  $\max\{\zeta_i, \xi_i\} < 1$ . Let the stopping time  $S$  be adapted to  $\mathcal{F}_t$  such that  $J_S := K_{1,S} + K_{2,S} \leq \eta$ , where  $\eta := \frac{2(C_1 + C_2)}{1 - \max\{\zeta_2, \dots, \zeta_{2m}, \xi_2, \dots, \xi_{2m}\}}$ . Furthermore, observe that for  $1 \leq i \leq 2m$

$$\begin{aligned} K_{1,t} K_{2,t+1} + K_{1,t+1} K_{2,t} &\geq \frac{u_{2i-2}^{t+1}}{u_{2i}^t} + \frac{u_{2i}^t}{u_{2i-2}^{t+1}} + \frac{u_{2i}^{t+1}}{u_{2i-2}^t} + \frac{u_{2i-2}^t}{u_{2i}^{t+1}} \\ (2.46) \quad &\geq \frac{u_{2i-2}^{t+1}}{u_{2i}^t} + \frac{u_{2i}^t}{u_{2i-2}^{t+1}} + 2 \end{aligned}$$

Therefore, substituting (2.46) in the definition of  $M_{t+1}$  and using the fact  $\frac{g_{2i-2}^{t+1}(v^t)}{v_{2i}^t} \leq \frac{u_{2i-2}^{t+1}}{u_{2i}^t}$  we get

$$\begin{aligned}
\mathbb{E}[R_{S+1}] &\leq \mathbb{E}\left[\left(\frac{1}{4}\dot{r}_{S+1,2m} + \frac{3}{4}\right)(R_S - 1)\right] + 1 \\
&\leq \mathbb{E}\left[\left(1 - \frac{1}{4}\left(\frac{1}{(1+2\mu_1\mu_2)M_{S+1}}\right)^{m-1}\left(\frac{1}{(1+2\mu_1\mu_2)\left(1+\frac{x}{v_2^S}\right)\left(1+\frac{u_2^S}{x}\right)}\right)\right)(R_S - 1)\right] + 1 \\
&\leq \mathbb{E}\left[\left(1 - \frac{1}{4}\left(\frac{1}{(1+2\mu_1\mu_2)(K_{1,S+1}K_{2,S} + K_{2,S+1}K_{1,S})}\right)^{m-1}\left(\frac{1}{(1+2\mu_1\mu_2)(1+xK_{1,S})\left(1+\frac{K_{2,S}}{x}\right)}\right)\right)(R_S - 1)\right] + 1 \\
&\leq \mathbb{E}\left[\left(1 - \frac{1}{4}\left(\frac{1}{(1+2\mu_1\mu_2)(K_{1,S+1}\eta + K_{2,S+1}\eta)}\right)^{m-1}\left(\frac{1}{(1+2\mu_1\mu_2)(1+x\eta)\left(1+\frac{\eta}{x}\right)}\right)\right)(R_S - 1)\right] + 1
\end{aligned}$$

Recall that  $K_{1,S+1} := \sum_{i=1}^m \frac{1}{u_{2i}^{S+1}}$  and  $K_{2,S+1} := \sum_{i=1}^m u_{2i}^{S+1}$ , and note that by (2.42) and (2.45) we get  $\mathbb{E}[K_{i,S+1} | \mathcal{F}_S] \leq \eta$  for  $i = 1, 2$ . Hence

$$\begin{aligned}
&\mathbb{E}\left[\left(1 - \frac{1}{4}\left(\frac{1}{(1+2\mu_1\mu_2)(K_{1,S+1}\eta + K_{2,S+1}\eta)}\right)^{m-1}\left(\frac{1}{(1+2\mu_1\mu_2)(1+x\eta)\left(1+\frac{\eta}{x}\right)}\right)\right)(R_S - 1)\right] + 1 \\
&\leq \mathbb{E}\left[\mathbb{E}\left[1 - \frac{1}{4}\left(\frac{1}{(1+2\mu_1\mu_2)(K_{1,S+1}\eta + K_{2,S+1}\eta)}\right)^{m-1}\left(\frac{1}{(1+2\mu_1\mu_2)(1+x\eta)\left(1+\frac{\eta}{x}\right)}\right) \middle| \mathcal{F}_S\right](R_S - 1)\right] + 1 \\
&\leq \left(1 - \frac{1}{4}\left(\frac{1}{(1+2\mu_1\mu_2)(\mathbb{E}[K_{1,S+1} | \mathcal{F}_S]\eta + \mathbb{E}[K_{2,S+1} | \mathcal{F}_S]\eta)}\right)^{m-1}\left(\frac{1}{(1+2\mu_1\mu_2)(1+x\eta)\left(1+\frac{\eta}{x}\right)}\right)\right) \mathbb{E}[(R_S - 1)] + 1 \\
(2.47) \quad &\leq r\mathbb{E}[(R_S - 1)] + 1
\end{aligned}$$

where  $r := 1 - \frac{1}{4}\left(\frac{1}{2\eta^2(1+2\mu_1\mu_2)}\right)^{m-1}\left(\frac{1}{(1+2\mu_1\mu_2)(1+x\eta)\left(1+\frac{\eta}{x}\right)}\right)$ . The transition to the last line of (2.47) is justified by Jensen's inequality. Also note that unlike in (2.9) where we concluded  $\mathbb{E}[R_{S+2}] \leq r\mathbb{E}[(R_S - 1)] + 1$ , we have  $\mathbb{E}[R_{S+1}] \leq r\mathbb{E}[(R_S - 1)] + 1$ . This is a result of directly using  $\{K_{1,S+1}, K_{2,S+1}\}$  without having to resort to a process like  $D_{S+1}$ , and implies that the factor  $\frac{1}{2}$  in the exponent of  $r$  in (2.9) can now be omitted. Therefore if we set  $J_t = K_{1,t} + K_{2,t}$  and define  $d'$  as before, then by the results of Lemma 11 and Lemma 12 and inequality (2.18), we get

$$(2.48) \quad \mathbb{E}[R_{t+1} | J_0 \leq \eta] \leq 1 + 3r^{\lfloor \frac{t}{d'} \rfloor} (R_0 - 1)$$

*Proof of Theorem 3.* The proof is identical to that of Theorem 1.  $\square$

We can confirm that for every  $n$  the condition  $\max\{\zeta_2, \dots, \zeta_{2m}, \xi_2, \dots, \xi_{2m}\} < 1$  is not vacuous, hence the previous results are applicable for certain parameter values. Observe first that if we set  $a_i := a$  for all  $i$ , then  $\zeta_2 = \frac{4a}{4(2a-1)}\left(1 + \sum_{j=2}^m \prod_{k=2}^j \frac{2a}{4(2a-1)}\right) < \frac{4a}{4(2a-1)} \cdot \frac{4(2a-1)}{2(3a-2)}$ , which is less than 1 whenever  $a > 2$ . Similarly, we conclude that whenever  $a > 5$ ,  $\zeta_{2m}, \xi_2$  and  $\xi_{2m}$  are all less than 1. Now for  $2 \leq i \leq m-1$ , we get

$$\begin{aligned}
(2.49) \quad \zeta_{2i} : &= \frac{a}{2(2a-1)} + \left(\frac{a}{(2a-1)} + \frac{a}{2(2a-1)} \frac{a}{2(2a-1)}\right) \left(1 + \sum_{j=i+1}^m \prod_{k=i+1}^j \frac{a}{2(2a-1)}\right) \\
&\leq q + (2q + q^2) \left(\frac{1 - q^{m+1}}{1 - q}\right)
\end{aligned}$$



where  $q = \frac{a}{2(2a-1)}$ . Since  $q \rightarrow \frac{1}{4}$  as  $a \rightarrow \infty$ , we conclude that  $\zeta_{2i} \rightarrow \frac{1}{4} + \frac{3}{4} \left(1 - \left(\frac{1}{4}\right)^{m+1}\right) < 1$ . Hence for  $a$  large enough,  $\zeta_{2i} < 1$  for all  $j$ , and a similar deduction follows for  $\xi_{2i}$ .

## APPENDIX

Before proceeding with bounding the two constants  $C_\pi$  and  $C_J$  appearing in the proof of Corollary 2, we will make the following definition used in deriving an upper bound for (2.53).

**Definition.** Let  $f \in L_1(\mathbb{R})$  be non-negative. Define  $\text{med}(f)$  to be the infimum over  $m \in \mathbb{R}$  such that

$$\int_{-\infty}^m f = \int_m^{\infty} f$$

**Lemma 15.** Suppose  $f(v) \in L_1(\mathbb{R})$  is non-negative, and  $g(v)$  is non-negative, monotone decreasing and  $fg(v) \in L_1(\mathbb{R})$ . Then  $\text{med}(fg) \leq \text{med}(f)$ .

*Proof.* Let  $m = \text{med}(f)$ . Then

$$\int_{-\infty}^m fg \geq g(m) \int_{-\infty}^m f = g(m) \int_m^{\infty} f \geq \int_m^{\infty} fg$$

Hence  $\text{med}(fg) \leq m \leq \text{med}(f)$ . □

**Corollary 16.** Suppose  $f(v) \in L_1(\mathbb{R}^+)$  is non-negative, and  $\sigma \geq 0$ . Then  $0 < y_1 \leq y_2$  implies  $\text{med}\left(\frac{f(v)}{(v+y_1)^\sigma}\right) \leq \text{med}\left(\frac{f(v)}{(v+y_2)^\sigma}\right)$ .

*Proof.* Since  $\left(\frac{v+y_2}{v+y_1}\right)^\sigma$  is monotone decreasing in  $v$  and  $\frac{f(v)}{(v+y_1)^\sigma} = \frac{f(v)}{(v+y_2)^\sigma} \left(\frac{v+y_2}{v+y_1}\right)^\sigma$ , the statement of the corollary is an immediate consequence of Lemma 15. □

*Proof of Corollary 2.* We can bound the ratio

$$(2.50) \quad \tilde{C}_\pi := \int \left( \frac{\max_i \{1, v_i\}}{\min_i \{1, v_i\}} \right) \left( \prod_{i=1}^4 v_i^{a_i + a_{i+1} - 1} \right) \exp \left( \sum_{i=1}^5 -v_i v_{i-1} \right) dv / C_g$$

appearing in the proof of Corollary 2 by first considering the simplification

$$(2.51) \quad \left( \frac{\max_i \{1, v_i\}}{\min_i \{1, v_i\}} \right) \leq \sum_{i=1}^4 \mathbf{1}_{\{v_i \geq 1\}} v_i + \sum_{i=1}^4 \mathbf{1}_{\{v_i \leq 1\}} \frac{1}{v_i} + \sum_{i \neq j; i, j \leq 4} \mathbf{1}_{\{v_i \geq 1, v_j \leq 1\}} \frac{v_i}{v_j}$$

It would therefore suffice to obtain an upper bound on the sum obtained by substituting (2.51) in the aforementioned ratio. Throughout this section we assume that the term  $N(w_1, w_2, w_3, w_4)$  (to be defined) is finite for values of  $(w_1, w_2, w_3, w_4)$  relevant to our computation. This will indeed be confirmed at the end.

Let  $N(w_1, w_2, w_3, w_4) := \int \left( \prod_{i=1}^4 v_i^{w_i - 1} \right) \exp \left( \sum_{i=1}^5 -v_i v_{i-1} \right)$ , and observe that by integrating w.r.t.  $v_1$  and  $v_4$  we get

$$(2.52) \quad N(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = \Gamma(\alpha_1) \Gamma(\alpha_4) \int \frac{v_2^{\alpha_2 - 1}}{(x + v_2)^{\alpha_1}} \frac{v_3^{\alpha_3 - 1}}{(b + v_3)^{\alpha_4}} \exp(-v_2 v_3) dv_2 dv_3$$

Hence

$$\begin{aligned} \frac{N(\alpha_1 + 1, \alpha_2, \alpha_3, \alpha_4)}{N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)} &= \alpha_1 \left( \int \frac{v_2^{\alpha_2 - 1}}{(x + v_2)^{\alpha_1 + 1}} \frac{v_3^{\alpha_3 - 1}}{(b + v_3)^{\alpha_4}} \exp(-v_2 v_3) dv_2 dv_3 \right) / \left( \int \frac{v_2^{\alpha_2 - 1}}{(x + v_2)^{\alpha_1}} \frac{v_3^{\alpha_3 - 1}}{(b + v_3)^{\alpha_4}} \exp(-v_2 v_3) dv_2 dv_3 \right) \\ &\leq \alpha_1 \left( \int \frac{1}{x} \frac{v_2^{\alpha_2 - 1}}{(x + v_2)^{\alpha_1}} \frac{v_3^{\alpha_3 - 1}}{(b + v_3)^{\alpha_4}} \exp(-v_2 v_3) dv_2 dv_3 \right) / \left( \int \frac{v_2^{\alpha_2 - 1}}{(x + v_2)^{\alpha_1}} \frac{v_3^{\alpha_3 - 1}}{(b + v_3)^{\alpha_4}} \exp(-v_2 v_3) dv_2 dv_3 \right) \\ &= \frac{\alpha_1}{x} \end{aligned}$$

By symmetry it follows immediately that  $N(\alpha_1, \alpha_2, \alpha_3, \alpha_4 + 1) / N(\alpha_1, \alpha_2, \alpha_3, \alpha_4) \leq \frac{\alpha_4}{b}$ .

We would now like to consider the ratio  $N_3(\alpha_1, \alpha_2, \alpha_3 + 1, \alpha_4) / N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ , where

$N_j(w_1, w_2, w_3, w_4) := \int \left( \mathbf{1}_{\{v_j \geq 1\}} v_j \prod_{i=1}^4 v_i^{w_i-1} \right) \exp\left(\sum_{i=1}^5 -v_i v_{i-1}\right) dv$ . Our goal is to arrive at an (good) upper bound for (2.50). When we substitute (2.51) in the integral in (2.50), we obtain (after moving the integral inside the summation) three summations of integrals, corresponding to the three summations in the right-hand side of (2.51). It follows that for the first summation it would be sufficient to consider the sum of ratios of this form since

$$\sum_{i=1}^4 \frac{N_i(\alpha_1, \dots, \alpha_i + 1, \dots, \alpha_4)}{N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)} = \int \left( \sum_{i=1}^4 \mathbf{1}_{\{v_i \geq 1\}} v_i \right) \left( \prod_{i=1}^4 v_i^{a_i + a_{i+1} - 1} \right) \exp\left(\sum_{i=1}^5 -v_i v_{i-1}\right) dv / C_g$$

Integrating first in  $v_2$  and  $v_4$ , we get

$$(2.53) \quad \frac{N_3(\alpha_1, \alpha_2, \alpha_3 + 1, \alpha_4)}{N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)} = \left( \int \frac{v_1^{\alpha_1-1}}{(v_1 + v_3)^{\alpha_2}} \frac{\mathbf{1}_{\{v_3 \geq 1\}} v_3^{\alpha_3}}{(v_3 + b)^{\alpha_4}} \exp(-xv_1) dv_1 dv_3 \right) / \left( \int \frac{v_1^{\alpha_1-1}}{(v_1 + v_3)^{\alpha_2}} \frac{v_3^{\alpha_3-1}}{(v_3 + b)^{\alpha_4}} \exp(-xv_1) dv_1 dv_3 \right)$$

Let

$$g(v_1) := \int \frac{\mathbf{1}_{\{v_3 \geq 1\}} v_3^{\alpha_3}}{(v_1 + v_3)^{\alpha_2} (v_3 + b)^{\alpha_4}} dv_3$$

Recall that  $\alpha_i = a_i + a_{i+1}$  and we assume apriori that  $a_i \geq 1$ , hence  $\alpha_2 + \alpha_4 - \alpha_3 = a_2 + a_5$ , and  $g$  is finite on  $\mathbb{R}^+$ . Furthermore, since  $g$  is monotone decreasing, it follows by Lemma 15 and from [6] that

$$\text{med}(g(v_1) v_1^{\alpha_1-1} \exp(-xv_1)) \leq \text{med}(v_1^{\alpha_1-1} \exp(-xv_1)) \leq \frac{\alpha_1}{x}$$

Thus

$$(2.54) \quad \int \frac{v_1^{\alpha_1-1}}{(v_1 + v_3)^{\alpha_2}} \frac{\mathbf{1}_{\{v_3 \geq 1\}} v_3^{\alpha_3}}{(v_3 + b)^{\alpha_4}} \exp(-xv_1) dv_1 dv_3 \leq 2 \int \frac{\mathbf{1}_{\{v_1 \leq \frac{\alpha_1}{x}\}} v_1^{\alpha_1-1}}{(v_1 + v_3)^{\alpha_2}} \frac{\mathbf{1}_{\{v_3 \geq 1\}} v_3^{\alpha_3}}{(v_3 + b)^{\alpha_4}} \exp(-xv_1) dv_1 dv_3$$

Then by applying Corollary 16 repeatedly, we can conclude (when considering the argument of  $\text{med}$  as a function of  $v_3$ ) that

$$(2.55) \quad \text{med} \left( \frac{\mathbf{1}_{\{v_1 \leq \frac{\alpha_1}{x}\}}}{(v_1 + v_3)^{\alpha_2}} \frac{\mathbf{1}_{\{v_3 \geq 1\}} v_3^{\alpha_3}}{(v_3 + b)^{\alpha_4}} \right) \leq \text{med} \left( \frac{\mathbf{1}_{\{v_3 \geq 1\}} v_3^{\alpha_3}}{\left(\frac{\alpha_1}{x} + v_3\right)^{\alpha_2} (v_3 + b)^{\alpha_4}} \right) \leq \text{med} \left( \frac{\mathbf{1}_{\{v_3 \geq 1\}} v_3^{\alpha_3}}{\left(v_3 + \frac{\alpha_1}{x} + b\right)^{\alpha_4 + \alpha_2}} \right)$$

We can bound the right-most term in (2.55) by the following method: note that  $\frac{\mathbf{1}_{\{v_3 \geq 1\}} v_3^{\alpha_3}}{\left(v_3 + \frac{\alpha_1}{x} + b\right)^{\alpha_4 + \alpha_2}} = \frac{\mathbf{1}_{\{v_3 \geq 1\}}}{v_3^{\alpha_4 + \alpha_2 - \alpha_3 - 0.5}} \frac{v_3^{\alpha_4 + \alpha_2 - 0.5}}{\left(v_3 + \frac{\alpha_1}{x} + b\right)^{\alpha_4 + \alpha_2}}$  and

$\frac{v_3^{\alpha_4 + \alpha_2 - 0.5}}{\left(v_3 + \frac{\alpha_1}{x} + b\right)^{\alpha_4 + \alpha_2}}$  is decreasing whenever  $v \geq 2 \left(\frac{\alpha_1}{x} + b\right) (\alpha_4 + \alpha_2 - 0.5)$ . Therefore, by Lemma 15

$$(2.56) \quad \begin{aligned} \text{med} \left( \frac{\mathbf{1}_{\{v_3 \geq 1\}} v_3^{\alpha_3}}{\left(v_3 + \frac{\alpha_1}{x} + b\right)^{\alpha_4 + \alpha_2}} \right) &\leq \text{med} \left( \frac{\mathbf{1}_{\{v_3 \geq \max\{1, 2(\frac{\alpha_1}{x} + b)(\alpha_4 + \alpha_2 - 0.5)\}\}} v_3^{\alpha_3}}{\left(v_3 + \frac{\alpha_1}{x} + b\right)^{\alpha_4 + \alpha_2}} \right) \\ &\leq \text{med} \left( \frac{\mathbf{1}_{\{v_3 \geq \max\{1, 2(\frac{\alpha_1}{x} + b)(\alpha_4 + \alpha_2 - 0.5)\}\}}}{v_3^{\alpha_4 + \alpha_2 - \alpha_3 - 0.5}} \right) \\ &= \max \left\{ 1, 2 \left( \frac{\alpha_1}{x} + b \right) (\alpha_4 + \alpha_2 - 0.5) \right\} 2^{\frac{1}{a_2 + a_5 - 1.5}} \end{aligned}$$

Setting  $m_1 = \max \{1, 2 \left( \frac{\alpha_1}{x} + b \right) (\alpha_4 + \alpha_2 - 0.5)\} 2^{\frac{1}{a_2+a_5-1.5}}$  and applying (2.54) and (2.55), we conclude

$$\begin{aligned}
 \frac{N_3(\alpha_1, \alpha_2, \alpha_3 + 1, \alpha_4)}{N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)} &= \left( \int \frac{v_1^{\alpha_1-1}}{(v_1 + v_3)^{\alpha_2}} \frac{\mathbf{1}_{\{v_3 \geq 1\}} v_3^{\alpha_3}}{(v_3 + b)^{\alpha_4}} \exp(-xv_1) \right) / \left( \int \frac{v_1^{\alpha_1-1}}{(v_1 + v_3)^{\alpha_2}} \frac{v_3^{\alpha_3-1}}{(v_3 + b)^{\alpha_4}} \exp(-xv_1) \right) \\
 &\leq \left( 2 \int \frac{\mathbf{1}_{\{v_1 \leq \frac{\alpha_1}{x}\}} v_1^{\alpha_1-1}}{(v_1 + v_3)^{\alpha_2}} \frac{\mathbf{1}_{\{v_3 \geq 1\}} v_3^{\alpha_3}}{(v_3 + b)^{\alpha_4}} \exp(-xv_1) \right) / \left( \int \frac{v_1^{\alpha_1-1}}{(v_1 + v_3)^{\alpha_2}} \frac{v_3^{\alpha_3-1}}{(v_3 + b)^{\alpha_4}} \exp(-xv_1) \right) \\
 &\leq \left( 4 \int \frac{\mathbf{1}_{\{v_1 \leq \frac{\alpha_1}{x}\}} v_1^{\alpha_1-1}}{(v_1 + v_3)^{\alpha_2}} \frac{\mathbf{1}_{\{v_3 \geq 1\}} \mathbf{1}_{\{v_3 \leq m_1\}} v_3^{\alpha_3}}{(v_3 + b)^{\alpha_4}} \exp(-xv_1) \right) / \left( \int \frac{v_1^{\alpha_1-1}}{(v_1 + v_3)^{\alpha_2}} \frac{v_3^{\alpha_3-1}}{(v_3 + b)^{\alpha_4}} \exp(-xv_1) \right) \\
 &\leq \left( 4m_1 \int \frac{\mathbf{1}_{\{v_1 \leq \frac{\alpha_1}{x}\}} v_1^{\alpha_1-1}}{(v_1 + v_3)^{\alpha_2}} \frac{\mathbf{1}_{\{v_3 \geq 1\}} \mathbf{1}_{\{v_3 \leq m_1\}} v_3^{\alpha_3-1}}{(v_3 + b)^{\alpha_4}} \exp(-xv_1) \right) / \left( \int \frac{v_1^{\alpha_1-1}}{(v_1 + v_3)^{\alpha_2}} \frac{v_3^{\alpha_3-1}}{(v_3 + b)^{\alpha_4}} \exp(-xv_1) \right) \\
 (2.57) \quad &\leq 4 \max \left\{ 1, 2 \left( \frac{\alpha_1}{x} + b \right) (\alpha_4 + \alpha_2 - 0.5) \right\} 2^{\frac{1}{a_2+a_5-1.5}}
 \end{aligned}$$

By the symmetry of  $N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$  we can also conclude that

$$N_2(\alpha_1, \alpha_2 + 1, \alpha_3, \alpha_4) / N(\alpha_1, \alpha_2, \alpha_3, \alpha_4) \leq 4 \max \left\{ 1, 2 \left( \frac{\alpha_4}{b} + x \right) (\alpha_3 + \alpha_1 - 0.5) \right\} 2^{\frac{1}{a_2+a_5-1.5}}$$

Observe next that

$$\begin{aligned}
 \frac{N(\alpha_1 - 1, \alpha_2, \alpha_3, \alpha_4)}{N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)} &= \frac{1}{\alpha_1 - 1} \left( \int \frac{v_2^{\alpha_2-1}}{(x + v_2)^{\alpha_1-1}} \frac{v_4^{\alpha_4-1}}{(v_2 + v_4)^{\alpha_3}} \exp(-bv_4) dv_2 dv_4 \right) / \left( \int \frac{v_2^{\alpha_2-1}}{(x + v_2)^{\alpha_1}} \frac{v_4^{\alpha_4-1}}{(v_2 + v_4)^{\alpha_3+1}} \exp(-bv_4) dv_2 dv_4 \right) \\
 &= \frac{1}{\alpha_1 - 1} \left( \int \frac{v_2^{\alpha_2-1} (x + v_2)}{(x + v_2)^{\alpha_1}} \frac{v_4^{\alpha_4-1}}{(v_2 + v_4)^{\alpha_3}} \exp(-bv_4) dv_2 dv_4 \right) / \left( \int \frac{v_2^{\alpha_2-1}}{(x + v_2)^{\alpha_1}} \frac{v_4^{\alpha_4-1}}{(v_2 + v_4)^{\alpha_3+1}} \exp(-bv_4) dv_2 dv_4 \right) \\
 &= \frac{1}{\alpha_1 - 1} \left( x + \frac{N(\alpha_1, \alpha_2 + 1, \alpha_3, \alpha_4)}{N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)} \right) \\
 &\leq \frac{1}{\alpha_1 - 1} \left( x + 1 + \frac{N_2(\alpha_1, \alpha_2 + 1, \alpha_3, \alpha_4)}{N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)} \right) \\
 &\leq \frac{1}{\alpha_1 - 1} \left( x + 1 + 4 \max \left\{ 1, 2 \left( \frac{\alpha_4}{b} + x \right) (\alpha_3 + \alpha_1 - 0.5) \right\} 2^{\frac{1}{a_2+a_5-1.5}} \right)
 \end{aligned}$$

The second-last inequality is a result of the fact that

$$N(\alpha_1, \alpha_2 + 1, \alpha_3, \alpha_4) = N_2(\alpha_1, \alpha_2 + 1, \alpha_3, \alpha_4) + \int \left( \mathbf{1}_{\{v_i < 1\}} v_i \prod_{i=1}^4 v_i^{\alpha_i-1} \right) \exp \left( \sum_{i=1}^5 -v_i v_{i-1} \right)$$

and the second term in the sum is less than or equal to the denominator  $N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ . Similarly we obtain

$$\frac{N(\alpha_1, \alpha_2, \alpha_3, \alpha_4 - 1)}{N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)} \leq \frac{1}{\alpha_4 - 1} \left( b + 1 + 4 \max \left\{ 1, 2 \left( \frac{\alpha_1}{x} + b \right) (\alpha_4 + \alpha_2 - 0.5) \right\} 2^{\frac{1}{a_2+a_5-1.5}} \right)$$

Proceeding in this manner we can also conclude that

$$(2.58) \quad \frac{N(\alpha_1, \alpha_2 - 1, \alpha_3, \alpha_4)}{N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)} = \frac{1}{\alpha_2 - 1} \left( \frac{N(\alpha_1 + 1, \alpha_2, \alpha_3, \alpha_4)}{N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)} + \frac{N(\alpha_1, \alpha_2, \alpha_3 + 1, \alpha_4)}{N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)} \right)$$

and

$$\frac{N(\alpha_1, \alpha_2, \alpha_3 - 1, \alpha_4)}{N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)} = \frac{1}{\alpha_3 - 1} \left( \frac{N(\alpha_1, \alpha_2, \alpha_3, \alpha_4 + 1)}{N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)} + \frac{N(\alpha_1, \alpha_2 + 1, \alpha_3, \alpha_4)}{N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)} \right)$$

Next we provide an upper bound for the terms pertaining to the sum  $\sum_{i \neq j; i, j \leq 4} \mathbf{1}_{\{v_i \geq 1, v_j \leq 1\}} \frac{v_i}{v_j}$  and its role in the ratio  $C_\pi$  in Corollary 2. Note that for the case  $i = 1, j = 2$  this is given by

$$(2.59) \quad \frac{N(\alpha_1 + 1, \alpha_2 - 1, \alpha_3, \alpha_4)}{N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)} = \frac{N(\alpha_1 + 1, \alpha_2 - 1, \alpha_3, \alpha_4)}{N(\alpha_1 + 1, \alpha_2, \alpha_3, \alpha_4)} \frac{N(\alpha_1 + 1, \alpha_2, \alpha_3, \alpha_4)}{N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)}$$

As was shown, the second term in the product on the right-hand side is bounded from above by  $\frac{\alpha_1}{x}$ , while the first term is of the same form as the term

$N(\alpha_1, \alpha_2 - 1, \alpha_3, \alpha_4) / N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$  and following an analogous derivation to (2.58) we can conclude that it is bounded from above by

$$\begin{aligned} \frac{N(\alpha_1 + 1, \alpha_2 - 1, \alpha_3, \alpha_4)}{N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)} &\leq \frac{\alpha_1}{x} \frac{1}{\alpha_2 - 1} \left( \frac{N(\alpha_1 + 2, \alpha_2, \alpha_3, \alpha_4)}{N(\alpha_1 + 1, \alpha_2, \alpha_3, \alpha_4)} + \frac{N(\alpha_1 + 1, \alpha_2, \alpha_3 + 1, \alpha_4)}{N(\alpha_1 + 1, \alpha_2, \alpha_3, \alpha_4)} \right) \\ &\leq \frac{\alpha_1}{x} \frac{1}{\alpha_2 - 1} \left( \frac{\alpha_1 + 1}{x} + 4 \max \left\{ 1, 2 \left( \frac{\alpha_1 + 1}{x} + b \right) (\alpha_4 + \alpha_2 - 0.5) \right\} 2^{\frac{1}{a_2 + a_5 - 1.5}} \right) \end{aligned}$$

A similar derivation follows for other values of  $i$  and  $j$ .

We can now summarise these results: let  $\varphi_1 = \frac{\alpha_1}{x}$ ,  $\varphi_2 = 4 \max \left\{ 1, 2 \left( \frac{\alpha_4}{b} + x \right) (\alpha_3 + \alpha_1 - 0.5) \right\} 2^{\frac{1}{a_2 + a_5 - 1.5}}$ ,  $\varphi_3 = 4 \max \left\{ 1, 2 \left( \frac{\alpha_1}{x} + b \right) (\alpha_4 + \alpha_2 - 0.5) \right\} 2^{\frac{1}{a_2 + a_5 - 1.5}}$  and  $\varphi_4 = \frac{\alpha_4}{b}$ . Also let  $\varphi_5 = \frac{1}{\alpha_1 - 1} (x + 1 + \varphi_2)$ ,  $\varphi_6 = \frac{1}{\alpha_2 - 1} (\varphi_1 + \varphi_3 + 1)$ ,  $\varphi_7 = \frac{1}{\alpha_3 - 1} (\varphi_4 + \varphi_2 + 1)$  and  $\varphi_8 = \frac{1}{\alpha_4 - 1} (b + 1 + \varphi_3)$ . Lastly let  $\varphi_9$  be same as  $\varphi_5$  but with every occurrence of  $\alpha_1$  replaced by  $\alpha_1 + 1$ , and similar definition follows for  $\varphi_{10}$ ,  $\varphi_{11}$  and  $\varphi_{12}$ . Then

$$(2.60) \quad \int \left( \frac{\max_i \{1, v_i\}}{\min_i \{1, v_i\}} \right) \left( \prod_{i=1}^4 v_i^{a_i + a_{i+1} - 1} \right) \exp \left( \sum_{i=1}^5 -v_i v_{i-1} \right) dv / C_g \leq \sum_{1 \leq i \leq 8} \varphi_i + \sum_{1 \leq i \leq 4, 9 \leq j \leq 12, j \neq i+8} \varphi_i \varphi_j$$

It remains to verify that  $N(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$  is in fact finite (the bounds in this section would then guarantee the finiteness of similar terms).

$$\begin{aligned} N(\alpha_1, \alpha_2, \alpha_3, \alpha_4) &= \Gamma(\alpha_2) \Gamma(\alpha_4) \int (\mathbf{1}_{\{v_1 + v_3 \leq 1\}} + \mathbf{1}_{\{v_1 + v_3 > 1\}}) \frac{v_1^{\alpha_1 - 1}}{(v_1 + v_3)^{\alpha_2}} \frac{v_3^{\alpha_3 - 1}}{(v_3 + b)^{\alpha_4}} \exp(-xv_1) \\ &= \Gamma(\alpha_2) \Gamma(\alpha_4) \left( \int \frac{\mathbf{1}_{\{v_1 + v_3 \leq 1\}} v_1^{\alpha_1 - 1} v_3^{\alpha_3 - 1} \exp(-xv_1)}{(v_1 + v_3)^{\alpha_2} (v_3 + b)^{\alpha_4}} + \int \left( \int \frac{\mathbf{1}_{\{v_1 + v_3 > 1\}} v_3^{\alpha_3 - 1}}{(v_1 + v_3)^{\alpha_2} (v_3 + b)^{\alpha_4}} dv_3 \right) v_1^{\alpha_1 - 1} \exp(-xv_1) dv_1 \right) \\ &\leq \Gamma(\alpha_2) \Gamma(\alpha_4) \left( \int \frac{\mathbf{1}_{\{v_1 + v_3 \leq 1\}} v_1^{\alpha_1 - 1} v_3^{\alpha_3 - 1} \exp(-xv_1)}{(v_1 + v_3)^{\alpha_2} (v_3 + b)^{\alpha_4}} + \frac{\Gamma(\alpha_1)}{x^{\alpha_1}} \left( \int_0^1 \frac{v_3^{\alpha_3 - 1} dv_3}{(v_3 + b)^{\alpha_4}} + \int_1^\infty \frac{dv_3}{v_3^{a_2 + a_5 + 1}} \right) \right) \\ &\leq \Gamma(\alpha_2) \Gamma(\alpha_4) \left( \int \frac{\mathbf{1}_{\{v_1 + v_3 \leq 1\}} \exp(-xv_1)}{(v_3 + b)^{\alpha_4}} + \frac{\Gamma(\alpha_1)}{x^{\alpha_1}} \left( \frac{1}{b^{\alpha_4}} + \int_1^\infty \frac{1}{v_3^{a_2 + a_5 + 1}} \right) \right) < \infty \end{aligned}$$

The last inequality follows from the fact that on  $\{v_1 + v_3 \leq 1\}$  we have  $(v_1 + v_3)^{\alpha_2} \geq (v_1 + v_3)^{\alpha_1 + \alpha_3 - 2} \geq v_1^{\alpha_1 - 1} v_3^{\alpha_3 - 1}$ .

The final bound is for  $C_J$ , which follows easily from the previous derivations. Recall that  $\mathcal{U}^0 = (1, 1, 1, 1)$  and  $\mathcal{V}^0 \sim \pi$ , while  $u_i^0 = \min \{\mathcal{U}_i^0, \mathcal{V}_i^0\}$  and  $v_i^0 = \max \{\mathcal{U}_i^0, \mathcal{V}_i^0\}$ , and note that

$$\begin{aligned}
 \mathbb{E}[J_0] &= \mathbb{E}[K_{1,0} + K_{2,0}] \\
 &= \mathbb{E}\left[u_2^0 + u_4^0 + \frac{u_3^0 + u_1^0 + b}{u_2^0(u_3^0 + u_1^0) + u_4^0(u_3^0 + b)}\right] \leq \mathbb{E}\left[u_2^0 + u_4^0 + \frac{1}{u_2^0} + \frac{1}{u_4^0}\right] \\
 &\leq \mathbb{E}\left[2 + \mathcal{V}_2^0 + \mathcal{V}_4^0 + 2 + \frac{1}{\mathcal{V}_2^0} + \frac{1}{\mathcal{V}_4^0}\right] \\
 (2.61) \quad &\leq 4 + \varphi_2 + \varphi_4 + \varphi_6 + \varphi_8
 \end{aligned}$$

Setting  $C_J$  equal to the right-hand side of the final inequality in (2.61) completes the proof of Corollary 2.  $\square$

*Acknowledgement.* The research of Neal Madras has been supported in part by a Discovery Grant from the Natural Science and Engineering Research Council of Canada.

#### REFERENCES

- [1] G.O. Roberts and J.S. Rosenthal (2004), General state space Markov chains and MCMC algorithms, Probability Surveys, Vol. 1, 20-71
- [2] T. Lindvall (1992), Lectures on the Coupling Method. Wiley&Sons, New York.
- [3] A. Gibbs (2004), Stochastic Models, Vol. 20, No. 4, 473-492
- [4] S. Geman and D. Geman (1984), Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images, IEEE Transactions on Pattern Analysis and Machine Intelligence (6), 721-741
- [5] L. Tierney (1994), Markov chains for exploring posterior distributions, Annals of Statistics, 22, 1701-1728
- [6] J. Chen and H. Rubin (1986), Statistics & Probability Letters 4, 281-283
- [7] M.F. Chen (1992), From Markov Chains to Non-Equilibrium Particle Systems, World Scientific
- [8] G.O. Roberts and J.S. Rosenthal (2002), One-shot coupling for certain stochastic recursive sequences, Stochastic Processes and their Applications, Vol. 99, 195-208